

FIGURE 1

CCAGGTCCAATGCCACCTCGTTCTATCGATTGAATTCCCCGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGTCCGCCAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGGCCTCAGGAGGTGCCTCAGGCCAGTGGCCTGAGGCCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGGCCACCATGCCACGCCGG
TCCAGCAGCATCAGCAGCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCACCCGGAGG
AGCAGCTCCTGCCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGA
AGGCCACCCGCCCTGGAGGCACAGGCCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCT
TCTGGTGTGGCAGTGGCGGACAGAGCACGCCCTACCGGCCGGCGTAGGGTGTGCT
GTCCCGGGCTCACGGGACCCCTGTCTCCAGTCGTTGTGCAGCGTGTGTACAGCCCTTCC
TCACCACCTGCGACGGGACCCGGCCTGCAGCACCTACCGAACCAATTATAGGACGCCCTAC
CGCCCGAGCCCTGGCCTGGCCCTGCCAGGCCCGCTACCGGTGCTGCCGGCTGGAAGAG
GACCAGCGGGCTTCCTGGGCCTGTGGAGCAGCAATATGCCAGGCCATGCCGGAACGGAG
GGAGCTGTGTCCAGCCTGGCGCTGCCGTGCCCTGCAGGATGGGGGTGACACTGCCAG
TCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGC
CGGCAGTTACTGGTGCCAGTGGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTGTG
TGCCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCGACAGGAGTGGACAGTGCAATGAAG
GAAGAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAACGCTGCAGCTGGT
GGCCCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCC
TCCTGGTGCACTCCTCCAGCAGCTGGCCGATCGACTCCCTGAGCGAGCAGATTCCCTC
CTGGAGGAGCAGCTGGGTCCCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGGCCAGG
CTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTC
CAGAAGCCACCTGGGGTGAUTGAGCGGAAGGCCAGGCAGGGCTTCCTCCTTCCCTC
CCCTCCCTGGGAGGGTCCCCAGACCCCTGGCATGGATGGCTGGGATTTTTGTGAAT
CCACCCCTGGTACCCCCACCCCTGGTACCCCAACGGCATCCCAAGGCCAGGTGGCCCTCA
GCTGAGGGAAGGTACGAGTTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGC
CGGAGGCTGGTGGGCCTCAGTGGGGCTGCTGCCGTGCCAGCCCCAGCACAATAAAATGAAA
CGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGACTCT
AGAGTCGACCTGCAGAAGCTTGGCCGCAATGCCCAACTTGTAAATTGAGCTTATAATGGT
TACAAAT

FIGURE 2

MTDSPPPGHPEEKATPPGGTGHEGLSGGAADVASGVGSGRHRARLPARPLGCVLSRAHDPV
SESFVQRVYQPFLTTCDGHRACSTYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGAC
GAAICQPPPCRNGGSCVQPGRCRCPAGWRGDTQSDVDECSARRGGCPQRCINTAGSYWCQCW
EGHSLSADGTLCPKGGPPRVAPNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLAS
QALEHGLPDPGSLLVHSFQQQLGRIDLSEQISFLEEQLGSCSCKKDS

Signal sequence:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 93-97, 270-274

N-myristoylation sites.

amino acids 19-25, 78-84, 97-103, 100-106, 103-109, 157-163,
191-197, 265-271

Amidation site.

amino acids 26-30

Aspartic acid and asparagine hydroxylation site.

amino acids 152-164

Cell attachment sequence.

amino acids 130-133

EGF-like domain cysteine pattern signature.

amino acids 123-135

FIGURE 3

CGCTCGCCCCGTGCCCTGCCCTCCCCGAGAGTCCCCCTCGCGGAGCAGATGTGTGGG
GTCAGCCCACGGCGGGACTATGGTGAAATTCCCGCGCTCACGCACTACTGGCCCTGATC
CGGTTCTTGGTGCCTGGCATCACCAACATAGCCATCGACTTCGGGGAGCAGGCCTTGAA
CCGGGGCATTGCTGCTCAAGGAGGATGCAGTCAGATGCTGGCCAGCTACGGCTGGCGT
ACTCCCTCATGAAGTTCTCACGGGTCCCATGAGTGACTTCAAAATGTGGGCTGGTGT
GTGAACAGCAAGAGAGACAGGACCAAGCCGTCTGTATGGTGGTGGCAGGGGCATCGC
TGCCGTCTTCACACACTGATAGCTTATAGTGAATTAGGATACTACATTATCAATAACTGC
ACCATGTGGACGAGTCGGTGGGGAGCAAGACGAGAAGGGCCTCCTGTACCTGCCGCCTT
CCTTCATGGACGCAATGGCATGGACCCATGCTGGCATTCTCTTAAACACAAATACAGTT
CCTGGTGGGATGTGCCTCAATCTCAGATGTAGCTCAGGGTTTTGTAGCCATTTGC
TTCACAGTCACCTGGAATGCCGGAGCCCTGCTCATCCGATCCTCTCCTGTACATGGC
GCACTTGTGCGCTGCACCACCCGTGCTGGCTACTACAAGAACATTACGACATCATCCC
TGACAGAAAGTGGCCGGAGCTGGGGAGATGCAACAATAAGAAAGATGCTGAGCTCTGGT
GGCCTTGGCTCTAATTCTGGCCACACAGAGAACAGTCCGTGCTGTATCCTGCTTCG
TCCCGGACCTGGTGGCAGTCTGCAGGCCACAGAGGAGTGGCATTGACAGCCACATA
CCCTGTGGTCACATGCCATACGGCTGGTACGGAAATCCGTGCTGTATCCTGCTTCG
ACAAGAATAACCCAGCAACAAACTGGTGAGCACGAGAACACAGTCACGGCAGCCCACATC
AAGAAGTTCACCTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTCTGTGATTTGGAC
ACCCAACGTGCTGAGAAAATCTGATAGACATCATGGAGTGGACTTGCCTTGAGAAC
TCTGTGTTGTCCTTGCAGTCTCTCCCTCCAGTTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAAAACCTTCGTCTTGCCCCAGCTCTGTGCG
GATCATCGTCTCATGCCAGCCTCGTGGCTACCTACGGGGTGACGGTGCACGGTGC
TGGCGTGGGCTCCCTCCGGGGCTTGTGGAGAACCTACCATGGTGCCTACGCTGCG
TGCTATGTCTACCGGAAGCAGAAAAAGAGATGGAGAACATGAGTCGGCACGGAGGGGAAGA
CTCTGCCATGACAGACATGCCCTCGACAGAGGAGGTGACAGACATCGTGGAAATGAGAGAGG
AGAATGAATAAGGCACGGGACGCCATGGGCACTGCAGGGACGGTCAGTCAGGATGACACTTC
GGCATCATCTTCCCTCTCCCATCGTATTTGTTCCCTTTTTGTTTTGGTAAT
GAAAGAGGCTTGATTAAAGGTTCTGTCAATTCTCTAGCATACTGGGTATGCTCACACT
GACGGGGGGACCTAGTGAATGGCTTACTGTTGCTATGAAAAAACGAAACAACTGAC
TTCATACCCCTGCCTCACGAAAACCAAAGACACAGTCGCTCACGGTTGACGTTGT
TCCTCCCTGGACAATCTCCTCTGGAACCAAGGACTGCAGCTGCCATCGGCCTCG
CACCGTGCACAGCAGGCCACAGACTCTCCTGTCCCCCTCATCGCTCTAAGAATCAACAGG
TTAAAACCGGCTTCTTGATTGCTTCCAGTCACATGCCGTACAAAGAGATGGAGCCC
CGGTGGCCTTAAATTCCCTCTGCCACGGAGTTGAAACCATCTACTCCACACATGCAG
GAGGCGGGTGGCACGCTGCAGCCGGAGTCCCCGTTCACACTGAGGAACGGAGACCTGTGAC
CACAGCAGGCTGACAGATGGACAGAACCTCCGTAGAAAGGTTGTTGAAATGCCCGGG
GGCAGCAAACGTACATGGTTGAATGATAGCATTCACTCTCGTTCTCCTAGATCTGAGCAA
GCTGTCAGTTCTCACCCCCACCGTGTATATACATGAGCTAACCTTTAAATTGTCAACAAA
GCGCATCTCCAGATTCCAGACCCGTGCCATGACTTTCTGAAGGCTTGCTTCCCTCG
CTTCTGAAGGTCGATTAGAGCGAGTCACATGGAGCATCCTAACCTTGCAATTAGTTT
TACAGTGAACTGAAGCTTAAGTCTCATCCAGCATTCTAAATGCCAGGTTGCTGTAGGTAAC
TTTGTAGATATATTACCTGGTTCTGCTATCCTAGTCATAACTCTCGGGTACAGGTAA
TTGAGAATGTACTACGGTACTTCCCTCCCACACCCATACGATAAACGAAAGACATTAAACG
ATACCAAGAGTCACATGTGGTCTCCCTGAAATAACGCATTGAAATCCATGCAGTGCAGTA
TATTCTAAGTTGGAAAGCAGGTTTTCCCTTAAAAAAATTATAGACACGGTTCACT
AAATTGATTTAGTCAGAATTCTAGACTGAAAGAACCTAAACAAAAAAATTAAAGATA
TAAATATATGCTGTATATGTTATGTAATTATTAGGCTATAACATACATTCCCTATTTCGC
ATTTCATAAAATGTCTCTAATACAAAAAA

FIGURE 4

MVKFPALTHYWPLIRFLVPLGITNIAIDFGEQALNRGIAAVKEDAVEMLASYGLAYSLMKFF
TGPMSDFKNVGLVFVNNSKDRTKAVLCMVVAGAIAAVFHTLIAYSDLGYYIINKLHHVDESV
GSKTRRAFLYLAAPFPMDAMAWTHAGILLKHKYSFLVGCAISDVIAQVVFVAILLHSHLEC
REPLLIPILSLYMGALVRCTTLCGYYKNIHDIIPDRSGPELGGDATIRKMLSFWWPLALIL
ATQRISRPIVNLFVSRDLGGSSAATEAVAILTATYPVGHMPYGWLTEIRAVYPAFDKNNPSN
KLVSTSNTVTAAHIKKFTFVCMALSLTLCFVMFWTPNVSEKILIDIIGVDFAFAELCVVPLR
IFSFFPVPVTVRALTGWLMTLKKTFLAPSSVLRIIVLIASLVLPYLGVHGATLGVGSL
AGFVGESTMVAIAACYVYRKQKKMENESATEGEDSAMTDMPPTEEVTDIVEMREENE

Transmembrane domains:

amino acids 86-106, 163-179, 191-205, 237-253, 327-343, 357-374,
408-423, 431-445

FIGURE 5

CCTGACAGAAGTGCCCGGAGCTGGGGAGATNCAACATTAAGAAGATGCTGAGCTTCTGGT
GCCNTTGGCTTAATTCTGCCACACAGAGAACAGTCGGCTATTGTCAACCTTTGTT
TCCCAGGACCTTGGTGGCAGTTCTGCAGCCACAGAGGCAGTGGCGATTTGACAGCCACATA
CCCTGTGGTCACATGCCATACGGCTGGTGACGAAATCCGTGCTGTATCCTGCTTCG
ACAAGAATAACCCAGCAACAAACTGGTGAGCACGAGAACACAGTCACGGGGCCCACATC
AAGAAGTTCACCTCGTCTGCATGGCTCTGTCACTCACGCTCTGTTCGTATGTTGGAC
ACCCAACGTGTCCTGNGAAAATCTTGATAGACATCATCGGAGTGGACTTGCCTTGAC
TCTGTGTTGTCCTTGGATCTCTCCTTCCAGTCCAGTCACAGTGAGGGCGCAT
CTCACCGGGTGGCTGATGACACTGAAGAACCTTCGTC

FIGURE 6

TGACGGAATCCCGGGCTGGGTATCCTGGTTNGACAAGATAAACCCCCAGCAANAAATTGGG
GAGCAGGGCAAAACAGTNACGGGCAGCCCACATCAAGAAGTTCACCTNGTTGNATGGNTC
TGTCAACTCACGCTNTGTTCGTATGTTGGACACCCAAAGTGTGAGAAAATTGAT
AGACATNATCGGAGTGGANTTGCCTTGCAGAANTTGNNGNTGTTGCCTTGCGGATTTCT
CCTTTTCCCAGTTCCAGTCACAGNGAGGGCGCATCTCACCGGGNGNTGATGACANTGAAG
AAAACCTTGTCTTGCCCCCAGCTNTTGGTGCAGTCATTGTCCTNATNGCCAGCCTTGT
GGTCCTACCTGGGGTGCACGGTGCACCGCTGGCGTGGGTCCCTCCTGGCGGGCA

FIGURE 7

TATTCCCAGTTCCGGTCACGGGGAGGGCGCATNTCACCGGGTGGCTGANGACACTGAAGAAA
ACCTTNGTCCTTGCCCCCAGNTTGTGNTGCGGATNATCGTCCTCATGCCAGCCTNGTGGT
CCTACCCCTACCTGGGGGTGCACGGTGAGAC

FIGURE 8

GCCTCGCTCCCTGCTCAGCTGCGCGTCTGCCTCTGGCTCTGCCCTGCATCCTGTGCAG
CTGCTGCCCGCCAGCCGCAACTCCACCGTGAGCCGCCATCTTACGTTCTCCTCTTCC
TGGGGGTGCTGGTGTCCATCATTATGCTGAGCCGGCGTGGAGAGTCAGCTCTACAAGCTG
CCCTGGGTGTGAGGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCACATGACTGTGG
CTCCCTGCTGGTACCGCGCTGTCTACCGATGTGCTCGCCACGGCGCCCTTCTTCTTCT
TCTTTTCACCCCTGCTCATGCTCTGCGTGAGCAGCAGCCGGACCCCCGGCTGCCATCCAG
AATGGGTTTGGTTCTTAAGTCCTGATCCTGGTGGGCTCACCGTGGTGCCTTACAT
CCCTGACGGCTCTTACCAACATCTGGTTCTACTTCGGCGTGTGGCTCCTCCTTCA
TCCTCATCCAGCTGGTGTGCTCATCGACTTGGCACTCCTGGAACCGGGTGGCTGGC
AAGGCCGAGGAGTGCGATTCCCGCCTGGTACCGAGGCCCTTCTTCTTCACTCTCCTTCA
CTACTTGCTGTCGATCGCGGCCGTGGCGCTGATGTTCATGTACTACACTGAGCCCAGGGCT
GCCACGAGGGCAAGGTCTTCATCAGCCTCAACCTCACCTCTGTGTCTCGTGTCCATCGCT
GCTGTCCTGCCCAAGGTCCAGGACGCCAGGCCAACTCAGGCTCTGCTGCAGGCCTCGTCAT
CACCCCTACACCATGTTGTCACCTGGTCAGGCCCTATCCAGTATCCCTGAACAGAAATGCA
ACCCCCATTGCCAACCCAGCTGGCAACGAGACAGTTGGCAGGCCCGAGGGCTATGAG
ACCCAGTGGTGGGATGCCCGAGCATTGTGGCCTCATCATCTCCTCCTGTGCACCCCTT
CATCAGTCTCGCTCCTCAGACCACCGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCC
CACCTATGCTAGACGCCACACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGCC
TTTGACAACGGAGCAGGACGGCGTCACCTACAGCTACTCCTCTTCACTTCTGCCTGGTGC
GGCCTCACTGCACGTCATGATGACGCTCACCAACTGGTACAAGCCGGTGAGACCCGGAAGA
TGATCAGCACGTGGACCGCCGTGGTGAAGATCTGTGCCAGCTGGCAGGGCTGCTCCTC
TACCTGTGGACCCCTGGTAGCCCCACTCCTCTGCCAACCGCGACTTCAGCTGAGGCAGCCT
CACAGCCTGCCATCTGGTGCCTCCTGCCACCTGGTGCCTCTGGCTCGGTGACAGCCAACCT
GCCCTCCCCACACCAATCAGCCAGGCTGAGCCCCACCCCTGCCAGCTCCAGGACCTG
CCCTGAGCCGGCCTTCTAGTCGTAGTGCCTCAGGGTCCGAGGAGCATCAGGCTCCTGCA
GAGCCCCATCCCCCGCCACACCCACACGGTGGAGCTGCCCTTCTCAGGCTCCAGGGAGCGGG
TGCCCATACTCAGCATCTCGGATGAAAGGGCTCCCTGTCCCTCAGGCTCCAGGGAGCGGG
CTGCTGGAGAGAGCGGGGAACCTCCACACAGTGGGGCATCCGGACTGAAGCCCTGGT
CCTGGTCACGTCCCCCAGGGACCCCTGCCCTCCTGGACTTCGTGCCTTACTGAGTCTCT
AAGACTTTTCTAATAACAAGCCAGTGCAGTGTAAAAAAA

FIGURE 9

MGACLGACSLSCASCLGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVE
SQLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFTLLMLCVSSRD
PRAAIQNGFWFKFLILVGLTVGAFYIPDGSFTNIWFYFGVVGGSFLFILQLVLLIDFAHSW
NQRWLGKAEECDSRAWYAGLFFFLLFYLLSIAAVALMFYYTEPSGCHEGKVFISLNLTFC
VCVSIAAVLPKVQDAQPNNSGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVA
GPEGYETQWWDAPSIVGLIIFLLCTLFISLRSSDHQVNNSLMQTEECPPMLDATQQQQQQVA
ACEGRAFDNEQDGVTYSYSFFHFCLVLASLHVMMTLTNWYKPGTRKMISTWTAVWVKICAS
WAGLLLLYLWTLVAPLLLNRDFS

Signal sequence:

amino acids 1-20

Transmembrane domains:

amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,
272-283, 324-340, 391-406, 428-444

FIGURE 10

GAGCGAGGCCGGGACTGAAGGTGTGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGTC
AAATGCACGGATTCTCACCTCGTACAGTTACGCTCTCCCGGGCACGTCCCGAGGACTTGA
AGTCCTGAGCGCTCAAGTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGGC
ACCGCGGAGCTTCTGTAGAGCATTGTGCCTATTCGGAGTCTTGCTGCCAAGCTG
TGACTGCCGATTCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTCCCTACGTCCCAGAGCCC
TATTACCGGAATCTGGATGGGACCGCCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAG
AATTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGG
TGTATGGGGAATACCAAGCTTTATTGCTAAACAACAATACATTGAGCAGAGCCAGGCA
GAAATTATCATAACCGGTTGATGCTGTGCAATCTGCACATCGTGCTGCCACACGAGGCTT
CATTGTTATGGCTGGCGCTGGGTTGGAGAACTGCACTGTTGTGACTATATTCAACACAG
TGAACACTAGTCTGAATGTATACGAAATAAAGATGCCTTAAGCCATTTGTAATTGCAGGA
GCTGTCACGGGAAGTCTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGCTGGCAT
AATTGGAGCCTGCTGGCACTCCTGTAGGAGGCCTGCTGATGGCATTCAAGTACGCTG
GTGAGACTGTTAGGAAAGAAAACAGAAGGATCGAAAGGCCTCCATGAGCTAAACTGGAA
GAGTGGAAAGGCAGACTACAAGTTACTGAGCACCTCCCTGAGAAAATTGAAAGTAGTTACG
GGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAAGCACTGCTAAACCTCCTAGAAACC
CTTCAGTAATAGATAAACAGACAAGGACTGAAAGTGCTCTGAAACTCACTGGAGA
GCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTGTCAGCCTGC
TGACAAATTAAAGTGTGGTACCTGTGGCAGTGGCTGCTTGTCTTTCTTTCTT
TTAACTAAGAATGGGCTGTTGACTCTCACTTACTTATCCTAAATTAAATACATACT
TATGTTGTATTAATCTATCAATATATGCATACGGATATATCCACCCACCTAGATTAA
GCAGTAAATAAACATTGCAAAAGATTAAAGTTGAATTACAGTT

FIGURE 11

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRILRELF
GKDEQQRISKDLANICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAEIYHNRFDAVQSAH
RAATRGFIRYGWRWGWRATAVFVTIFNTVNTSLNVYRNKDALSHFVIAGAVT GSLFRINVGLR
GLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDRKALHELKLEEWKGRLQVTEHLPE
KIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

Important Features:

Signal Peptide:

amino acids 1-24

Transmembrane domains:

amino acids 76-96 and 171-195

N-glycosylation site:

amino acids 153-156

FIGURE 12

CGGAAGTCCCTGAGGAGCGTCAGAAGCGGCTTCCCTACGTCCCAGAGCCCTATTACCCGGA
ATCTGGATGGGACCGCTCCGGAGCTGTTGGCAAAGATGAACAGCAGAGAATTCAAAGGA
CCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGAA
TACCAGCTTTATTGATGCTGTGCAATCTGCACATCGTGTGCCACACGAGGCTTCATTGTTCATG
AACCGGTTGATGCTGTGCAATCTGCACATCGTGTGCCACACGAGGCTTCATTGTTCATG
GCTGGCGCCGAACC

FIGURE 13

TCAAGTTGTCCGTAGGTCGAGAGAAGGCCATGGAGGTGCCGCCACCGCACCGGGAGCTT
TTTCTGTAGAGCATTGTGCCTATTCCCCGAGTTTGCTGCCGAAGCTGTGACTGCCGAT
TCGGAAGTCCTTGAGGAGCGTCAGAACGGGCTTCCCTACGTCCCAGAGCCCTATTACCGGA
ATTGGATGGGACCGCCTCCGGGAGCTGTTGGCAAAGATGAACAGCAGAGAATTCAAAGG
ACCTTGCTGATATNTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGA
ATACCAGCTTTATTCATGNTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTATNA
TAACC

FIGURE 14

GAGCCGCCGCCGCCGCCGCCGCCACTGCAGCCCCAGGCCGCCACCGTCT
GCGTTGCTGCCCGCCTGGGCCAGGCCAAAGGCAAGGACAAAGCAGCTGTCAAGGAACCT
CCGCCGGAGTCGAATTACGTGCAGCTGCCGAACACAGGTTCAAGATGGTTGCCGGG
GCTTCGCGTGTCCAAGAACTGCCTGTGCCCTCAACCTGCTTACACCTGGTTAGTCTG
CTGCTAATTGGAATTGCTGCGTGGGCATTGGCTCAGGCTGATTCCAGTCTCGAGTGGT
CGGCGTGGTCATTGCAGTGGCATCTCTTGTGATTGCTTAGTGGCTGATTGGAG
CTGTAAAACATCATCAGGTGTTGCTATTTTATATGATTATTCTGTTACTGTATTATT
GTTCAGTTTCTGTATCTGCGCTTGTAGCCCTGAACCAGGAGAACAGGGTCAGCTCT
GGAGGTTGGTGGAACAAATACGGCAAGTGCTCGAAATGACATCCAGAGAAACTAAACTGCT
GTGGGTTCCGAAGTGTAAACCAAATGACACCTGTCTGGTAGCTGTGTTAAAAGTGACCAC
TCGTGCTGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTGAGATTGTTGG
TGGCATTGGCCTGTTCTCAGTTACAGAGATCCTGGGTGTTGGCTGACCTACAGATA
GGAACCAGAAAGACCCCCCGCGAATCCTAGTGCATTCTTGATGAGAAAACAAGGAAGAT
TTCCTTCGTATTATGATCTTGTCACTTCTGTAATTTCTGTTAAGCTCCATTGCCAGT
TTAAGGAAGGAAACACTATCTGGAAAAGTACCTTATTGATAGTGGAAATTATATTTACT
CTATGTTCTACATGTTTCTTCCGTTGCTGAAAAATATTGAAACTTGTGGTCTC
TGAAGCTCGGTGGCACCTGGAATTACTGTATTCACTGTGGCAGTCCACTGTGGCCTT
TCTTAGCATTACCTGCAGAAAAACTTGTATGGTACCACTGTGTTGGTATATGGTGA
TCTGAACGTACATCTCACTGGTATAATTATATGTAGCACTGTGCTGTAGATAGTCCCTAC
TGGAAAAAGAGTGGAAATTATTAAAATCAGAAAGTATGAGATCCTGTTATGTTAAGGGAAA
TCCAAATTCCAATTGGTCTTTAGGAAAGATTGTTGGTAAAAAGTGTAGTA
TAAAAATGATAATTACTGTAGTCTTATGATTACACCAATGTATTCTAGAAATAGTTAT
GTCTTAGGAAATTGTGGTTAATTGACTTTACAGGTAAGTGCAAAGGAGAACGTGGTT
CATGAAATGTTCTAATGTATAATAACATTACCTCAGCCTCCATCAGAACGGAGTTT
TGAGTAATCAGGAAGTATATCTATGATCTGATATTGTTATAATAATTGAAGTCTAA
AAGACTGCATTAAACAAAGTTAGTATTACGTTGGCCACGTAGCAAAAGATATTG
ATTATCTTAAAAATTGTTAAATACCGTTTACGAAATTCTCAGTATTGTAACAGCAACTT
GTCAAACCTAAGCATATTGAATATGATCTCCATAATTGAAATTGAAATCGTATTGTGTG
GCTCTGTATATTCTGTTAAAAATTAAAGGACAGAACCTTCTTGTATGCATTTGA
ATTAAAAGAAAGTAATGGAAG

FIGURE 15

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979
><subunit 1 of 1, 204 aa, 1 stop
><MW: 22147, pI: 8.37, NX(S/T): 3
MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVVIAVGIFLFLIALV
GLIGAVKHHQVLLFFYMIILLLVFIVQFSVSCACLAQNQEQQGQLLEVGNNTASARNDIQR
NLNCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRFVGGIGLFFSFTEILGVWL
TYRYRNQKDPRANPSAFL
```

Signal Peptide:

amino acids 1-34

Transmembrane domains:

amino acids 47-63, 72-95 and 162-182

FIGURE 16

TGATTGGAGCTGTAAAAAANTCTTCAGGTGTTGNATTTTTATGATTATTCTGTAANT
TGTATTTATTGTCAGTTNTGTATCTGCGCTTGTAGCCNTGAACCAGGAGCAACAGG
GTCAGNTNTGGAGGTTGGTGGAAACAATACGGCAAGTGCTCGAAATGACATCCAGAGAAAT
NTAAACTGCTGTGGGTTCCGAAGTGTAAACCAAATGACACCTGTNTGGCTAGCTGTGTTAA
AAGTGACCACTNGTGCTGCCATGTGCTCCAATCATAGGAGAATATGCTGGAGAGGTTTGA
GATTGTTGGTGGCATTGGCCTGTTNTCAGTTTACAGAGATCCTGGGTGTTGGCTGACC
TACAGATAACAGGAACCAAG

FIGURE 17

AATCCCAAATTCCCCAATTTTTGGNCTTTAGGGAAAGATGTGTTGTGGTAAAAAGTGT
TAGTATAAAAATGATAATTTACTTGTAGTCTTTATGATTACACCAATGTATTCTAGAATAG
TTATGTCTTAGGAAATTGTGGTTAATTTTGACTTTACAGGTAAGTGCAAAGGAGAAGTG
GTTTCATGAAATGTTCTAATGTATAATAACATTACCTTCAGCCTCCCATCAGAATGGAACG
AGTTTGAGTAATCCAGGAAGTATATCTATATGATCTTGATATTGTTTATATAATTTGAAG
TCTAAAAGACTGCATTTAAACAAGTTAGTATTAATGCGTTGGCCCACGTAGCAAAAGAT
ATTTGATTATCTTAAAAATTGTTAAATACCGTTTCATGAAAGTTCTCAGTATTGTAACAGC
AACTTGTCAAACCTAACGATATTGAATATGATCTCCATAATTGAAATTGAAATCGTATT
GTGTGGAGGAAATGGCAATCTTATGTGTGCTGAAGGACACAGTAAGAGCACCAAGTTGTGCC
CCACTTGC

FIGURE 18

ATGATTATTCTGTTACTGTATTATTGTTCAGTTTATGGTATCTTGCCTGTTAGCCC
CTGAAACCAGGAGCAACAGGGNNCAGCTTCTGGAGGTTGGCAACAAATCACGGCCAAG
TGACTCCGCAAATGACATCCCAGAGAAACTCTAAACTGCTGTGGTTCCGAAGTGTAAACCC
AAATGACACACTGTCTGGCTNGCTGTGTTAAAGTGACCACTCGTGCTGCCATGTGCTCCAA
TCATAGGAGAATATGC

卷之三

FIGURE 19

CAGTCACCATGAAGCTGGCTGTGCTCATGGCCTGGCCCTACCTTCCCTGGTGTG
CTCTGGGTGGCCAGATGCTACTGGCTGCCAGTTGAGACGCTGAGTGTGAGGGACCTGT
CTGCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTGACTGATGCAAGGGAAGCTGGCT
TCCAGGTCAAGGCCTACACTTCAGTGAACCCCTCCACCTGATTGTGTCTATGACTGGCTG
ATCCTCCAAGGTCCAGCCAAGCCAGTTTGAAGGGGACCTGCTGGTCTGCGCTGCCAGGC
CTGGCAAGACTGCCACTGACTCAGGTGACCTCTACCGAGATGGCTCAGCTCTGGTCCCC
CCGGGCCTAACAGGAAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCAC
TGCAGTGGCATCTCCAGGCCCTGGTCTGGATCCCAGAAACAGCATCTGTTGTGGCTAT
CACAGTCCAAGAACTGTTCCAGGCCAATTCTCAGAGCTGTACCCCTCAGCTGAACCCAAG
CAGGAAGCCCCATGACCTGAGTTGTCAAGACAAAGTTGCCCTGCAGAGGTCACTGCCGC
CTCCTCTCTCCTCTAACAGGATGGAAGGATAGTCAAAGCAGGGGCTCTCCTCAGAATT
CCAGATCCCCACAGCTCAGAAGATCACTCCGGTCATACTGGTGTGAGGCAGCCACTGAGG
ACAACCAAGTTGGAAACAGAGCCCCAGCTAGAGATCAGAGTGCAGGGTGTCCAGCTCT
GCTGCACCTCCCACATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAACTGCTCCTGA
GGAGGCCCTGGCCTCTGCCCTCGCCGCCAACCCATCTCTGAGGATCCAGGCTTTCTT
CTCCTCTGGGATGCCAGATCCTCATCTGTATCACCAGATGGCCTTCTCAAACACATG
CAGGATGTGAGAGTCCTCCTCGGTACCTGCTCATGGAGTTGAGGAATTATCTGCCACCA
GAAGCCTGGGACCACAAAGGCTACTGCTGAATAGAAGTAAACAGTTCATCCATGATCTCACT
TAACCACCCAAATAATCTGATTCTTATTTCTCTTCTGTCACATATGCATAAGTA
CTTTACAAGTTGTCCAGTGTGTTAGAATAATGTAGTTAGGTGAGTGTAAATAAATT
ATATAAAAGTGAGAATTAGAGTTAGCTATAATTGTGTATTCTCTCTAACACACAGAATT
TGCTGTCTAGATCAGGAATTCTATCTGTTATATCGACCAAGATGTTGTGATTAAAGAGAA
CTAATGGAAGTGGATTGAATACAGCAGTCTCAACTGGGGCAATTGCCCCCAGAGGACA
TTGGCAATGTTGGAGACATTGTCATTATACTTGGGGGTTGGGGATGGTGGATGT
GTGTCTACTGGCATCCAGTAAATAGAAGCCAGGGTGCCGCTAACACATCCTATAATGCACAG
GGCAGTACCCACAACGAAAAATAATCTGGCCAAAATGTCAGTTGTTACTGAGTTGAGAAA
CCCCAGCCTAATGAAACCCTAGGTGTTGGCTCTGGAATGGACTTTGTCCTCTAATT
TATCTCTTCCAGCCTCATTCACTATTCTACTGACATACCAGTCTTAGCTGGTGTATG
GTCTGTTCTTAGTTCTAGTTGTATCCCTCAAAAGCCATTATGTTGAAATCCTAATCCCC
AAGGTGATGGCATTAAAGAAGTGGCCTTGGGAAGTGTAGATCAGGAGTGCAGAGCCCTC
ATGATTAGGATTAGTGCCCTATTTAAAAGGCCAGAGAGCTAACTCACCCCTCCACCAT
ATGAGGACGTGGCAAGAAGATGACATGTATGAGAACCAAAACAGCTGTGCCAAACACCG
ACTCTGTCGTGCCTTGATCTGAACCTCCAGCCTCCAGAACTATGAGAAATAAAATTCTGG
TTGTTGTAGCCTAA

FIGURE 20

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLGCVLMWALYLSLGVLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQV
KAYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVRLCQAWQDWPLTQVTFYRDGSALGPPGP
NREFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGS
PMTLSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHGSYWCCEAATEDNQ
VWKQSPQLEIRVQGASSSAAPPTLNPAQKSAAPGTAPEEEAPGPLPPPPTPSSEDPGFSSPL
GMPDPHLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE
```

Signal sequence:

amino acids 1-17

Leucine zipper pattern sequence:

amino acids 12-33

Protein kinase C phosphorylation site:

amino acids 353-355

FIGURE 21

CCACCGCGTCCGCCACCGTCCGCCACGGTCCGCCACCGTCCGGCCACCGAGTT
TGAGCCTTTGGTAGCAGGAGGCTGGAAGAAAGGACAGAAGTAGCTCTGGCTGTGATGGGG
ATCTTACTGGCCTGCTACTCCTGGGCACCTAACAGTGGACACTTATGGCCGTCCCACCTCCT
GGAAGTGCCAGAGAGTGTAAACAGGACCTGGAAAGGGGATGTGAATCTCCCTGCACCTATG
ACCCCTGCAAGGCTACACCCAAGTCTGGTAAGTGGCTGGTACAACGTGGCTCAGACCC
GTCACCATCTTCTACGTGACTCTTCTGGAGACCATATCCAGCAGGAAAGTACCAAGGGCCG
CCTGCATGTGAGCCACAAGGTTCCAGGAGATGTATCCCTCCAATTGAGCACCCCTGGAGATGG
ATGACCGGAGCCACTACACGTGTGAAGTCACCTGGCAGACTCCTGATGGCAACCAAGTCGTG
AGAGATAAGATTACTGAGCTCCGTGTCAGAAACTCTCTGTCTCCAAGCCCACAGTGACAAC
TGGCAGCGGTTATGGCTTCACGGTCCCCAGGGATGAGGATTAGCCTCAATGCCAGGCTC
GGGTTCTCCCTCATCAGTTATATTGGTATAAGCAACAGACTAATAACCAGGAACCCATC
AAAGTAGCAACCTAAGTACCTACTCTCAAGCCTGCGGTGATAGCCGACTCAGGCTCCTA
TTCTGCACTGCCAAGGGCCAGGGTGGCTCTGAGCAGCACAGCAGACATTGTGAAGTTGTGG
TCAAAGACTCCTCAAAGCTACTCAAGACCAAGACTGAGGCACCTACAACCATGACATAACCC
TTGAAAGCAACATCTACAGTGAAGCAGTCTGGACTGGACCACTGACATGGATGGCTACCT
TGGAGAGACCAGTGTGGGCCAGGAAAGAGCCTGCGCTGTCTTGCCATCATCCTCATCATCT
CCTTGTGCTGTATGGGGTTTTACCATGGCCTATATCATGCTCTGTGCGGAAGACATCCAA
CAAGAGCATGTCTACGAAGCAGCCAGGAAAGAAAGTCTCTCCTCTCCATTGGACCCGT
CCCTGCCCTCAATTGATTACTGGCAGGAAATGGAGGAAGGGGGGTGTGGCACAGACCC
AATCCTAAGGCCGGAGCCTCAGGGTCAGGACATAGCTGCCCTCCCTCTCAGGCACCTT
CTGAGGTTGTTGGCCCTCTGAACACAAAGGATAATTAGATCCATGCCCTCTGCTTCC
AGAATCCCTGGTGGTAGGATCTGATAATTAAATTGGCAAGAATTGAGGCAGAAGGGTGGGA
AACCAAGGACCACAGCCCCAAGTCCCTCTTATGGGTGGTGGCTCTGGGCCATAGGGCACA
TGCCAGAGAGGCCAACGACTCTGGAGAAACCATGAGGGTGGCATCTCGCAAGTGGCTGCT
CCAGTGATGAGCCAACCTCCCAGAATCTGGCAACAACTACTCTGATGAGCCCTGCATAGGA
CAGGAGTACCAAGATCATGCCAGATCAATGGCAACTACGCCGCTGCTGGACACAGTTCC
TCTGGATTATGAGTTCTGGCCACTGAGGGCAAAAGTGTCTGTTAAAATGCCCTATTAGGC
CAGGATCTGCTGACATAATTGCTAGTCAGTCCTGCTGATGCCCTCTCCCTGCT
ACCTCTCTGGATAGCCAAAGTGTCCGCTACCAACACTGGAGCCGCTGGAGTCAC
GGCTTGGCCCTGGAATTGCCAGATGCATCTCAAGTAAGCCAGCTGCTGGATTGGCTCTGG
GCCCTCTAGTATCTGCCGGGGCTCTGGTACTCCTCTAAATACCAGAGGGAAAGATG
CCCATAGCACTAGGACTGGTCATCATGCCCTACAGACACTATTCAACTTGGCATCTGCCA
CCAGAAGACCCGAGGGAGGCTCAGCTGCCAGCTCAGAGGACCAAGCTATATCCAGGATCAT
TTCTCTTCTCAGGGCCAGACAGCTTTAATTGAAATTGTTATTTCACAGGCCAGGGTTCA
GTTCTGCTCCACTATAAGTCTAATGTTCTGACTCTCCTGGTGTCAATAATATCTA
ATCATAACAGC

FIGURE 22

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45416
><subunit 1 of 1, 321 aa, 1 stop
><MW: 35544, pI: 8.51, NX(S/T): 0
MGILLGLLLGHLTVDTYGRPILEVPESVTGPWKGDVNLPCTYDPLQGYTQVLVKWLVQRGS
DPVTIFLRDSSGDHIQQAKYQGRLHVSHKVPGDVSLQLSTLEMDDRSHYTCEVTWQTPDGNQ
VVRDKITELRVQKLSVSKPTVTTGSGYGYFTVPQGMRISLQCQARGSPPISYIWYKQQTNQ
PIKVATLSTLLFKPAVIADSGSYFCTAKGQVGSEQHSDIVKFVVKDSSKLLKTKTEAPTTMT
YPLKATSTVKQSWDWTMDGYLGETSAGPGKSLPVFAIILIIISLCCMVFTMAYIMLCRKT
SQQEHVYEAAR
```

Signal Sequence:

amino acids 1-19

Glycosaminoglycan attachment site:

amino acids 149-152

Transmembrane domain:

amino acids 282-300

FIGURE 23

GCGCCGGGAGCCCATCTGCCCCAGGGGCACGGGGCGGGGCGCTCCGCCGGCACAT
GGCTGCAGCCACCTCGCGCGACCCCGAGGCGCCAGCTGCCCGAGGTCCGTGGA
GGCGCCGGCCGCCCCGGAGCCAAGCAGCAACTGAGCGGGGAAGGCCCGGTCCGGGATC
GGGATGTCCCTCCTCTCTCTCTCTGCTAGTTCTACTATGTTGAACTTGGGACTCA
CACTGAGATCAAGAGAGTGGCAGAGGAAAAGGTCACTTGCCTGCCACCATCAACTGGGC
TTCCAGAAAAAGACACTCTGGATATTGAATGGCTGCTACCGATAATGAAGGGAACCAAAA
GTGGTGTACACTTACTCCAGTCGTATGTCATAAATACTGACTGAGGAACAGAACAGGGCG
AGTGGCTTGTCTCCAATTCTGGCAGGAGATGCCCTTGCAAGATTGAACCTCTGAAGC
CCAGTGATGAGGGCCGGTACACCTGTAAGGTTAAGAATTAGGGCCTACGTGTGGAGCCAT
GTCATCTAAAAGTCTTAGTGAGACCATCCAAGCCAAGTGTGAGTTGGAAGGAGAGCTGAC
AGAAGGAAGTGAACCTGACTTGCAGTGTGAGTCATCCTCTGGCACAGAGCCCATTGTGTATT
ACTGGCAGCGAATCCGAGAGAAGAGGGAGAGGATGAACGTCTGCCCTCCAAATCTAGGATT
GACTACAACCACCCCTGGACGAGTTCTGCTGCAGAATCTTACCATGCTCTACTCTGGACTGTA
CCAGTGCACAGCAGGCAACGAAGCTGGGAAGGAAAGCTGTGTGGTGAGTAACTGTACAGT
ATGTACAAAGCATGGCATGGTGCAGGAGCAGTGACAGGCATAGTGGCTGGAGCCCTGCTG
ATTTCTCTGGTGTGGCTGCTAATCCGAGGAAAGACAAAGAAAGATATGAGGAAGAAGA
GAGACCTAATGAAATTGAGAAGATGCTGAAGCTCCAAAGCCGTCTGTGAAACCCAGCT
CCTCTCCTCAGGCTCTGGAGCTCACGCTCTGGTTCTCCTCCACTCGCTCCACAGCAAAT
AGTGCCTCAGCAGCCAGCGAACACTGTCACACTGACGCAGCACCCAGCCAGGGCTGCCAC
CCAGGCATAACGCCTAGTGGGCCAGAGGTGAGAGGTTCTGAACCAAAGAAAGTCCACCATG
CTAATCTGACCAAAGCAGAAACCACACCCAGCATGATCCCCAGCCAGAGCAGGCCTCCAA
ACGGTCT**GA**ATTACAATGGACTTGACTCCCACGCTTCTAGGAGTCAGGGTCTTGGACTC
TTCTCGTCATTGGAGCTCAAGTCACCAGCCACACAACCAGATGAGAGGTCATCTAAGTAGCA
GTGAGCATTGACGGAACAGATTAGATGAGCATTTCTTATACAATACCAAACAAGCAA
AGGATGTAAGCTGATTCTGTAAAAAGGCATCTTATTGTGCCTTAGACCAAGAGTAAGGG
AAAGCAGGAGTCCAAATCTATTGTTGACCAGGACCTGTGGTGAGAAGGTTGGGAAAGGTG
AGGTGAATATACTCTAAACTTTAATGTTGAGATATTGTATCAGTGTCTTGATTACAATT
TTCAAGAGGAATGGGATGCTGTTGTAATTTCTATGCATTCTGCAAACCTATTGGATT
ATTAGTTATTAGACAGTCAGCAGAACCCACAGCCTTATTACACCTGTCTACACCATGTAC
TGAGCTAACCACTCTAAGAAACTCCAAAAAGGAAACATGTGTCTCTATTCTGACTTAAC
TTCATTGTCTAAGGTTGGATTTAATTCAGGGGAGTTGAAATAGTGGGAGATGGAGA
AGAGTGAATGAGTTCTCCACTCTATACTAATCTCACTATTGTATTGAGCCAAAATAAC
TATGAAAGGAGACAAAATTGTGACAAAGGATTGTGAAGAGCTTCCATCTCATGATGTT
ATGAGGATTGTTGACAAACATTAGAAATATATAATGGAGCAATTGTGGATTCTCCCTCAAAT
CAGATGCCCTCTAAGGACTTCTGCTAGATATTCTGGAAAGGAGAAAATACAACATGTCATT
TATCAACGTCCCTAGAAAGAATTCTCTAGAGAAAAGGATCTAGGAATGCTGAAAGGATTA
CCCAACATACCAATTATAGTCTCTTCTGAGAAAATGTGAAACCCAGAATTGCAAGACTGG
GTGGACTAGAAAGGGAGATTAGATCAGTTCTCTTAATATGTCAAGGAAGGTAGCCGGCA
TGGTGCAGGCACCTGTAGGAAAATCCAGCAGGTGGAGGTTGCAAGTGAAGCCGAGATTATGCC
ATTGCACTCCAGCCTGGGTGACAGAGCGGGACTCCGTCTC

FIGURE 24

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45419
><subunit 1 of 1, 373 aa, 1 stop
><MW: 41281, pI: 8.33, NX(S/T): 3
MSLLLLLLVSYYVGTGTHTEIKRVAEEKVTLPCHHQLGLPEKDTLDIEWLTDNEGNQKV
VITYSSRHVYNNLTEEQKGRVAFASNFLAGDASLQIEPLKPSDEGRYTCKVKNSGRYVWSHV
ILKVLVRPSKPCKELEGELTEGSDLTLQCESSSGTEPIVYYWQRIREKEGEDERLPPKSRID
YNHPGRVLLQNLTMYSGLYQCTAGNEAGKESCVRVTVQYVQSIGMVAGAVTGIVAGALLI
FLLVWLLIRRKDKEYEEEERPNEIREDAEAPKARLVKPSSSSGSRSSRGSSSTRSTANS
ASRSQRTLSTDAAQPGLATQAYSLVGPEVRGSEPKVHHANLTKAETTPSMIPSQSRAFQTV
```

Signal sequence:

amino acids 1-16

Transmembrane domain:

amino acids 232-251

FIGURE 25

FIGURE 26

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594
><subunit 1 of 1, 655 aa, 1 stop
><MW: 71845, pI: 8.22, NX(S/T): 8
MGTSPSSSTALASCSRIARRATATMIAGSLLLLGFLSTTTAQPEQKASNLIGTYRHVDATG
QVLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFRHENGIEKCHDCSQPCPWPMIEKLPCA
ALTDRECTCPPGMFQSNATCAPHTVCVGWGVRKKGTEVEDVRCKQCARGTFSDVPSSVMKC
KAYTDCLSQNLLVVIKPGTKETDNVCGTLPSFSSSTSPSPGTIAIFPRPEHMETHEVPSSTYVP
KGMNSTESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHIL
KLLPSMEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVVIVVCSIRK
SSRTLKKGPRQDPSAIVEKAGLKKSMPTQNRKWIYYCNGHgidILKLVAAQVGSQWKDIY
QFLCNASEREVAAFSNGYTADHERAYAALQHWTIRGPEASLAQLISALRQHRRNDVVEKIRG
LMEDTTQLETDKLALPMSPSPLSPSPIPSPNAKLENSALLTVEPSPQDKNGFFVDESEPLL
RCDSTSSGSSALSRNGSFITKEKKDTVLQVRLDPCLQPIFDDMLHFLNPEELRVIEEIPQ
AEDKLDRRLFEIIGVKSQEASQTL LDSVYSHLPDLL
```

Signal sequence:

amino acids 1-41

Transmembrane domain:

amino acids 350-370

FIGURE 27

ATGGGAAGCCAGTAACACTGTGGCCTACTATCTCTCCGTGGTGCACATCTACATTTTGGGA
CTCGGGATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCTGAAATAGTCAC
CATGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTCTCATTCCGATCGCTTTGGCC
TTGATGATTGAAAATAAGTCCTGTTGACCCAGATGCAGATGCTGTTGCTGACAGATCCTG
TCACTGCTGCCATTGAAGTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGC
ACTGCCATTGGTCTGGGCATCCACTCGACTGCTCAGGAAGTACAGATGTCGCTCATCCT
TTAAGTGTATCGAGCTGATAGCTCGATGTGACGGAGTCTGGATTGCAAAGACGGGAGGAC
GAGTACCGCTGTGTCCGGTGGTGGTCAGAATGCCGTGCTCCAGGTGTTACAGCTGCTTC
GTGGAAGACCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAAC
TGGGTTTCCAAGCTATGTGAGTTCAAGATAACCTCAGAGTGGAGCTCGCTGGAGGGCAGTTC
CGGGAGGAGTTGTGTCATCGATCACCTCTGCCAGATGACAAGGTGACTGCATTACACCA
CTCAGTATATGTGAGGGAGGGATGTGCCTCTGCCACGTGGTTACCTGCAGTGCACAGCCT
GTGGTCATAGAAGGGCTACAGCTCACGCATCGTGGGTGAAACATGTCCTGCTCTCGCAG
TGGCCCTGGCAGGCCAGCCTCAGTCCAGGGTACACACCTGTGCGGGGCTCTGTCATCAC
GCCCTGTGGATCATCACTGCTGCACACTGTGTTATGACTTGTACCTCCCCAAGTCATGGA
CCATCCAGGTGGTCTAGTTCCCTGTTGGACAATCCAGCCCCATCCACCTGGTGGAGAAG
ATTGTCTACCACAGCAAGTACAAGCCAAAGAGGCTGGCAATGACATGCCCTATGAAGCT
GGCCGGGCCACTCACGTTCAATGAAATGATCCAGCCTGTGCGCTGCCAACCTGAAGAGA
ACTTCCCCGATGGAAAAGTGTGCTGGACGTCAGGATGGGGGCCACAGAGGATGGAGGTGAC
GCCTCCCTGTCCTGAACCACGCGGCCGTCCTTGATTCCAACAAAGATCTGCAACCACAG
GGACGTGTACGGTGGCATCATCTCCCCCTCCATGCTCTGCCGGCTACCTGACGGTGGCG
TGGACAGCTGCCAGGGGACAGGGGGGCCCTGGTGTCAAGAGAGGAGGCTGTGGAAG
TTAGTGGGAGCGACCAGCTTGGCATCGCTGCCAGGGTGAACAAGCCTGGGTGTACAC
CCGTGTCACCTCCCTGGACTGGATCCACGAGCAGATGGAGAGAGACCTAAAAACTGAA
GAGGAAGGGACAAGTAGCCACCTGAGTCCCTGAGGTGATGAAGACAGCCGATCCTCCCC
GGACTCCCGTGTAGGAACCTGCACACGAGCAGACACCCCTGGAGCTGAGTCCGGCACCA
GTAGCAGGCCGAAAGAGGCACCCCTCCATCTGATTCCAGCACAACCTCAAGCTGCTTTT
GTTTTTGTGTTTTGAGGTGGAGTCTCGCTCTGTTGCCAGGCTGGAGTGCAGTGGCGAAA
TCCCTGCTCACTGCAGCCTCCGCTTCCCTGGTCAAGCGATTCTCTGCCCTCAGCTCCCC
GTAGCTGGGACCACAGGTGCCGCCACCACACCCAACTAATTTGTATTTAGTAGAGAC
AGGGTTCACCATGTTGCCAGGCTGCTCTCAAACCCCTGACCTCAAATGATGTGCCTGCTT
CAGCCTCCCACAGTGTGGATTACAGGCATGGGCCACCACGCCTAGCCTCACGCTCCTTTC
TGATCTCACTAAGAACAAAAGAAGCAGCAACTGCAAGGGCGCCTTCCCACTGGTCCAT
CTGGTTTCTCTCCAGGGTCTTGCAAAATTCTGACGAGATAAGCAGTTATGTGACCTCACG
TGCAAAGCCACCAACAGCCACTCAGAAAAGACGCACAGCCAGAAGTGCAGAACTGCAGTC
ACTGCACGTTTCATCTCTAGGGACCAGAACCAACCCACCCCTTCTACTTCCAAGACTTAT
TTTCACATGTGGGAGGTTAATCTAGGAATGACTCGTTAAGGCCTATTTCATGATTCTT
TGTAGCATTGGTGTGACTGAGTATTATTGTCCTTGATTCCAATAATGTTCCCTTCCCT
CATTGTCGGCGTGTGCGTGGACTGGTGAATCAAATCATCCACTGAAA

FIGURE 28

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234
><subunit 1 of 1, 453 aa, 1 stop
><MW: 49334, pI: 6.32, NX(S/T): 1
MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILA
LAIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVLQVFTAAS
WKTMCSDDWKGHYANVACAQLGFPSYVSSDNLRVSSLEQFREFVSIDHLLPDDKVTALHH
SVYVREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVIT
PLWIITAHCVDLYLPKSWTIQVGLVSILLDNPAPSHLVEKIVYHSKYKPKRNLGNDIALMKL
AGPLTFNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPLISNKICNHR
DVYGGIISPSMLCAGYLGGVDSQGDGGPLVCQERRLWKLVGATSGIGCAEVNKPGVYT
RVTSLDWIHEQMERDLKT
```

Signal Peptide:

amino acids 1-20

Transmembrane domain:

amino acids 240-284

FIGURE 29

CCCACCGCGTCCGTCTAGTCCCCGGCCAACTCGACAGTTGCTCATTATTGCAACGGTCAAGGCTGGCTTGT
GCCAGAACGGCGCGCGCGCACGCACGCACACACACACGGGGAAACTTTTAAAATGAAAGGCTAGAAGA
GCTCAGCGGCGCGCGGGCGCTGCGGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA
CGCCCGGCCCCGGCTCGGCGCCCGCGTGGGATGGTGCAGCGCTCGCCGGGGCCGAGAGCTGCTGCACTGAAG
GCCGGCGACGTGGCAGCGCGCCCGTCCCGTGTCCCCCGCCCGCCCTCCTGCTCGCCCTGGCCGGTGTCT
GCTCGCGCCCTGCGAGGCCCAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTCAGTGCCTCTGT
TCGGAGTGGGACTCTGGATCCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAAGTGTGAATATTGACT
ACAACGGGAAAGCAAAGAAGTGTACATAAAATCTGAAAGAAATGAAGGCTCATTGCCAGCAGTTCACGAAAC
CCACTATCTGCAAGACGGTACTGATGTCCTCGCTCGAAATTACACGGGTCAGTGTACTACCATGGACATGT
ACGGGGATATTCTGATTCAAGCAGTCAGTCTCAGCACGTGTTGGTCTCAGGGGACTTATTGTGTTGAAATGA
AAGCTATGTCTTAAACCAATGAAAAGTGAACCAACAGATAACAAACTCTTCCCAGCGAAGAAGCTGAAAAGCGT
CCGGGGATCATGTGGATCACATACAACACACAAACCTCGCTGCAAAGAATGTGTTCCACCAACCCCTCAGAC
ATGGGCAAGAAGGCATAAAAGAGAGACCTCAAGGCAACTAAGTATGTGGAGCTGGTGTGACGGCAGACAAACCG
AGAGTTTCAAGAGGCAAGGAAAGATCTGAAAAGTTAACGAGCATTAAATAGAGATTGCTAATCACGTTGACAA
GTTTACAGACCACTGAACATTGGATCGTGTGGTAGGCAGGGTGGAAAGTGTGGAATGACATGGACAAATGCTCTGT
AAGTCAGGACCCATTCAACCAGCCTCCATGAATTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA
TGACAATGCGCAGCTGTCAAGTGGGTTTATTCCAAGGGACCCATCGGATGGCCCAATCATGAGCATGTG
CACGGCAGACCCAGCTGGGGAAATTGTCATGGACCAATTCAAGACAATCCCTTGGTGCAGCCGTGACCCCTGGCACA
TGAGCTGGGCCAACATTTCGGGATGAACCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA
AGGAGGCTGCATCATGAACCGCTCCACCGGGTACCCATTCCATGGTCTCAGCAGTGTGACGGAGACTT
GGAGACCCAGCCTGGAGAAAGGAATGGGGTGTGCTGTTAACCTGCCGGAAGTCAGGGAGCTTCCGGGCA
GAAGTGTGGGAAACAGATTGTGGAAGAAGGAGGGAGTGTGACTGTGGGAGCCAGAGGAATGTATGAATCGCTG
CTGCAATGCCACCCACTGTACCTGAAAGCCGACGCTGTGCGCACATGGGCTGTGCTGTGAAAGACTGCCAGCT
GAAGCCTGCAGGAACAGCGTCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTCTGACACAGGGGCCAGCCC
TCACTGCCAGCAATGTGACCTGACGATGGGACTCATGTGAGGATGTGACGGCTACTGCTACAATGGCAT
CTGCCAGACTCACGAGCAGCAGTGTGACGCTTGGGACCAAGGTGCTAACACTGCCCTGGGATCTGCTTGA
GAGAGTCAAATTCTGCAAGGTGATCCTTATGGCAACTGTGGCAAAGTCTGAAAGAGTTCTTGCCTGGGAAATGCGAGAT
GAGAGATGCTAAATGTGGAAAATCCAGTGTCAAGGAGGTGCAAGCCGGAGTCATTGGTACCAATGCCCTTC
CATAGAAAACAACATCCCTCTGCAGCAAGGAGGCCGGATTCTGTGCCGGGGACCCACGTGTACTTGGGCGATGA
CATGCCGGACCCAGGGCTTGTGCTGAGGACAAAGTGTGCAAGATGGAAAATCTGCTGAATCGTCAATGTCA
AAATATTAGTGTCTTGGGTTACGAGTGTGCAATGCACTGGCAGGGGTGTGCAACAAACAGGAAGAA
CTGCCACTGCGAGGCCACTGGCACCTCCCTCTGTGACAAGTTGGCTTGGAGGAAGCACAGACAGGGGCC
CATCCGCAAGCAGAAGCAAGGAGGAAGCTGCAAGGTCAAACAGGGAGCGCGGCCAGGGCAGGAGCCGGTGGG
ATCGCAGGAGCATGCGTCACTGCCTCACTGACACTCATTGAGCCCTCCATGACATGGAGACCGTGAACAGTG
CTGCTGAGAGGAGGTCAAGCGTCCCCAAGGCCCTGTGACTGGCAGCATTGACTCTGTGGCTTGCCTCGT
TCCATGACAACAGACACAACAGTTCTGGGCTCAGGAGGGAAAGTCAGCCCTACCGCACGTCTGCAGAAA
CACTGCAAGGAGGGCAGCGACTCTCTGGTTGAGCTTCTGCTAAACATGGACATGCTCAGTGTGCTCCTGAG
AGAGTAGCAGGTTACCACTCTGGCAGGGCCCAGGCCCTGCAGCAAGGAGGAAGGAGACTCAAAGTCTGGCTTTC
ACTGACGCCCTCCACAGCAGTGGGGAGAAGCAAGGGTTGGGCCAGTGTCCCCCTTCCCACTGACACCTCAGCCT
TGGCAGGCCCTGATGACTGGCTCTGGCTGCAACTTAATGCTCTGATGCTTGCATTTAGCATTATATGAAAAT
AGCAGGGTTTTAGTTAAATTATCAGAGACCCCTGCCACCCATTCCATCTCCAGCAAGCAAACGTGAATGGCAA
TGAAAACAACGTGAGAAGGAAGGTAGGAGAAAGGGCGGTGAACCTGGCTTTGCTGGACATGCGTGACCGAGC
AGTACTCAGGTTGAGGGTTGCAAGAACCCACAGAGTCACCAACCCCTTCAATTAAACAAGTAAGAA
TGTAAAAAGTGAACAAATGTAAGAGCCTAACTCCATCCCCGTGGCATTACTGCATAAAATAGAGTGCATT
GAAAT

FIGURE 30

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624
><subunit 1 of 1, 735 aa, 1 stop
><MW: 80177, pI: 7.08, NX(S/T): 5
MAARPLPVSPARALLALAGALLAPCEARGVSLWNQGRADEVVSASVRSGDLWIPVKSFDISK
NHPEVLNIRLQRESKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVRG
YSDSAVSLSTCSGLRGLIVFENESYVLEPMKSATNRYKLFPAKKLKSVRGSCGSHHNTPNLA
AKNVFPPPSQTWARRHKRETLKATKYVELVIVADNREFQRQGKDLKVKQRLIEIANHVDKF
YRPLNIRIVLVGVEVWNDMDKCSVSQDPFTSLHEFLDWRKMLLPRKSHDNAQLVSGVYFQG
TTIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNFGMNHDTLRGCSCQMAVEK
GGCIMNASTGYPFPVMFSSCSRKDLETSLEKGMGVCLFNLPLEVRESFGGQKCGNRFVEEGEE
CDCGEPEECMNRC-CNATTCTLKPD-CAHGLC-CEDCQLK-PAGTACRDSSNSCDLPEFCTGAS
PHCPANVYLHDGHSCQDV-DGYCYNGICQ-THEQQCVTLWGP-GAKPAPGICF-ERVNSAGDPYGN
CGKVS-KSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIE-TNIPLQQGGRILCRGTHVYLG
DDMPDPGLVLAGTKCADGKICLNRCQNI-SVFGVHECAMQCHGRGVCNNRKNC-CEAHWAPP
FCDKFGFGSTDGPIRQAEARQEAESNRERGQGQEPVGSQEHA-STASLT
```

Signal peptide:

amino acids 1-28

FIGURE 31

TCCCAAGGCTTCTTGGATGGCAGATGATTNTGGGGTTGCATTGTTCCCTGACAACGAAA
ACAAAACAGTTGGGGTTCAGGAGGGGAANTCCAGCCTACCCAGGAAGTTGCAGAAACA
GTGCAAGGAAGGGCAGGANTCCTGGTTGAGNTTTGNTAAAACATGGACATGNTTCAGTG
CTGCTCNTGAGAGAGTAGCAGGTTACCACCTTGCGAGGCCAGCCCTGCAGCAAGGAGGA
AGAGGACTCAAAAGTTGGCCTTCACTGAGCCTCACAGCAGTGGGGAGAAGCAAGGGTT
GGGCCAGTGTCCCCTTCCCCAGTGACACCTCAGCCTGGCAGCCCTGATAACTGGTNTNT
GGCTGCAANTTAATGCTNTGATATGGCTTTAGCATTATTATGAAAATAGCAGGGTTTT
AGTTTTAATTATCAGAGACCCGCCACCCATTCCATNTCCATCCAAG

FIGURE 32

Top 2000000

CATCCTGCAACATGGTGAACACCACGCCGGCTAATTTGTTGTTAGAGATGGGA
TTTCACCGTGTAGCCAGGATTGTCTCAATCTGACCTCATGATCTGCCCGCCTGGCCTCCC
AAAGTGCTGGGATTACAGGCAGTGCAACCACACCCGGCCACAAACTTTAAGAAGTTAAT
GAAACCATACTTTACATTTAATGACAGGAAAATGCTCACAATAATTGTTAACCCAAAA
TTCTGGATACAAAAGTACAATCTTACTGTGTAATACATGTATATGTAATGAAAATA
TACCAAATATCAATAACTTATCTGGTAAAAACCTCTCATACCCGTGCTAACAA
CTTTAACAAAAAATTCATCACTTTAAGAATCAAGAAAATTCATGAAGGTCAATGGG
ACAGAAAAAAACCAAGGGAAAATCACGCCACTTGGGAAAAAGATTGAAATCTGCCT
TTTATAGATTGTAATTAATAAGGTCCAGGCTTCTAAGCAACTAAATGTTGTTGCA
AACAAAGTACTTGTCTGGATGTAGGAGGAAAGGGAGTGATGTCAGGCCATTATGATGCC
TTGAATATAAGACCCTACTGCTATCTCCCTGCACCAGCCAGGAGCCACCCATCCTCCAGC
ACACTGAGCAGCAAGCTGGACACACGGCACACTGATCCAAATGGTAAAGGGATGGTGGCGA
TGCTCATTCTGGGCTGCTACTTCTGGCGCTGCTCCTACCCGTGCAGGTTCTCATTGTT
CCTTTAACCAAGTATGCCCGAAGCTACTGCAGCCGAAACCACAAAGCCCTCCAACAGTGCC
ACAGCCTACAGCCGGTCTCCTGTGGCTTGCTTGCCTTCTACATCTTACCAATGGTAAAGGGATGGTGGCGA
CAGGTCAAGAAACAGCTACAGTTCTCAACCCATAACTAAAACGAATCCAATGGTGCCT
AGAAGTTCAATGTGGCAAGGAAAAAAACCAAGGTCTTACATCAAATCTACTAATTCAC
ATTAACAGAGAAACGCTTGAGAGTCTCAAACGGACTGGACTGGTTAAAGAGCATCTGAAGGATT
GACTAGATGATAATGCCGTACTCCAGTACTTGGGAGGCCTAGGCCGGGATCACCTG
AGGTCAAGGAGTTGAGACTAACCTGGCAAAATGGTGAACCCCCATCTGTACTAAAATACA
AATATTGACTGGCGTGGTGGTGAGTGCCTGTGATCCAGCTACTCAGGTGGCTGAAGCAGG
ACAATCACTGAACTCAGGAGGCAGAGGTTGCAGTGAGCTGAGATCGCGCTACTGCACTCTA
GCCTAGCCTGGCAACAGAGTGAGACTTCGTCTCAAAAAAAAGCCAAGTGCAGTGGCT
CACGCCGTAAATCCCGCACTTGGGAGGCCAGGTGGCGGATCAGGAGTCAGGAGATCA
AGACCACCTGGCTAATACAGTGAAACCCGTCTACTAAAATACAAAAATTAGCCGG
GATGGTGGCAGGCACCTGGAGTCCAGCTACTCGGGAGGCTGAGGCAGGAGAATAGCGTGAA
CTCAGGAGGCAGGAGCTTGAGTGCAGTGAGCCGAGATTGCGCTACTGCACTCCAGCCTGGCGACAG
CGCGAGACTCCGTCTCAAAAAAAAAAAAAAA

FIGURE 33

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48309
><subunit 1 of 1, 67 aa, 1 stop
><MW: 6981, pI: 7.47, NX(S/T): 0
MGKGMVAMLILGLLLALLLPVQVSSFVPLTSMPEATAAETTKPSNSALQPTAGLLVVLLAL
LHLYH
```

Signal peptide:

amino acids 15-27

.....

FIGURE 34

GGCGCGCGAGAGCGCGCCAGCCCCCGCGATGCCGCGCCCAGGACGCCTCTCCGCTGCTGGCCCGC
CGGGCCCTGACTGCGCTGCTGCTGCTGGCCATGGCGGGCGCTGGGCGCCAGGGCCAG
AGGCGGCGGCGGCGGCGGCGGACGGGCCCCCGCGGAGACGGCGAGGACGGACAGGACCCGACAGCAAGCACC
TGTACACGGCCGACATGTTACGCACGGGATCCAGAGCGCCGCACTTCGTATGTTCTCGCCCTGGTGTG
GACACTGCCAGCGGCTGAGCGACTTGAATGACCTGGGAGACAAATAACAGCATGGAAGATGCCAAAGTCT
ATGTTGCTAAAGTGGACTGACGGCCACTCCGACGCTGCTCCGCCCAGGGGTCGAGGATACCCACCTTAA
AGCTTCAAGCCAGGCAAGAAGCTGTAAGTACCAAGGGCTCTCGGGACTTCCAGACACTGGAAAAGTGGATGC
TGCAGACACTGAACGAGGAGGGCAGTGACACCAGAGCCGGAAGTGGAAACGCCAGTGCCTGGAGCTCAAGCAAG
GGCTGTATGAGCTCTGAGCAAGCAGGACTGAGCTGACGGTCAAGGGACCACTTATCAAGTTCTCGCTC
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCAAACCTGGAGCAGCTGGCTCTGGGCTTGAACATTCCGAAACTG
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCGGAAACAGGGTTGTGGCTACTGAGGGAGTACG
TTCTCTGGTCCAGATGGAAAAGGGTGGATCAAGGGAAAGGGGATTGGAGTCAGTGAGGGAGTACG
TGGAGTCGCACTGAGCGCACAGAGACTGGAGCGACGGAGACCGTACGCCCTAGAGGCCCGGTGCTGGCAG
CTGAGCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAATAACTTCGATGACACCATTGAGAAGGAA
TAACCTTCATCAAGTTTATGCTCCATGGTGTGGTCAAGTGAAGACTCTGGCTCCTACTTGGGAGGAACCTCT
AAAAGGAATTCCCTGGTCTGGGGGGTCAAGATGCCAGTAGACTGCACTGCTGAACGGAAATATCTGAGCA
AGTATTGCGTACGGAGGCTACCCCACGTTATTGCTTTCCAGGGAGGAAGAAAGTCAGTGAGCACAGTGGAGGCA
GAGACCTTGACTCGTTACACCGCTTGCTCTGAGCCAAGCGAAAGACGAACCTT~~AGGAACACAGTTGGAGGTAC~~
CTCTCCTGCCAGCTCCGACCCCTGCGTTAGGAGTTCACTGCCCACAGGGCCTGGGTTCCAGTGGTGGCT
GTTCAAGAAAGCAGAACATAACTAAGCGTGAGGTATCTCTTGTTGTGTGTTTCCAAGCCAACACACTCTACAG
ATTCTTTATTAAGTTAAGTTCTAAGTAAATGTGTAACTCATGGTCACTGTTAAACATTTCAGTGGCATA
TATCCCCTTGACCTTCTCTGATGAAATTACATGGTTCTTGAGACTAAAATAGCGTTGAGGGAAATGAAA
TTGCTGGACTATTGTTGTCCTGAGTTGAGTGAATTGGTGAAGAACACATCCAAAGCATACTTACCTGC
CCACCGAGTTCTGAAAGGTGGCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTACAGTTGACTCAATAC
TGTGTTGGTCCCGTAGCATGGAGCAGATTGAAATGCCAAAACCACACCTCTGGAAGATACTTACGGCCGCTGC
TGGAGCTTCTGCTGTGAATACTCTCAGTGTGAGAGGTTAGCCGTGATGAAAGCAGCGTTACTCTGACC
GTGCTGAGTAAGAGAATGTCATGCCATAACTTATGTCGATACTTGTCTCATGTTAAACACGAAGGCCAGGCCACATGA
CCTTCTGTTCTCACGGGGTGAACATGTCATTAGGCTCATGTTAAACACGAAGGCCAGGCCACATGA
TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATTCCACGGCTCATGTTCACTGTTGACTCT
AAGATTCCAGTTGATTGTCACCTGGGGTACAAGACCGAGACAGGCTTCCAGGCCGGTATCCAGGGAGG
TCTGCAGCCCTGCTGAAGGGCCCTAACAGAGTTCTAGAGTTCTGATTCTCAGTAGTCCTTTAGAGG
CTTGCTATACTTGGTCTGCTTCAAGGAGGTGACCTCTAAATGTAAGAAATGGGATGCAATTGATCTCAAGAC
CAAAGACAGATGTCAGTGGCTGCTCTGGCCCTGGTGTGACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCT
ACTCATGCTGCTTGTGATTAACACCTCTATCTCCCTGGAAATAAGCACACAGGTTAAGCTCTAAGATA
GATAGGTGTTGCTTTTACCATCGAGCTACTTCCATAATAACCACTTGCATCCAACACTCTCACCCACCT
CCCATACGCAAGGGGATGTTGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA
CTGTCGTGTCAGGGCAGAAGATAACAGCAGCATCTCGACCGAGCCCTGCTTAAAGGAATCTTATTAAATCAG
TATGGTTCAAGATAATTCTTTTAAAGGGAAACCTCTTAGAGAAAGCACAACGTCAAGAGTCCTGTACA
CACAACCTCAGCTTGTGATCACGAGTCCTGTTCAAGAAAATCAAAGTGGTACAATTGTTGTTACACTAT
GATACTTTCTAAATAACTCTTTTTTAA

FIGURE 35

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0
MPARPGRLLPPLLARPAALTALLLLLLGHGGGGRWGARAQEAAAAADGPPAADGEDGQDPHS
KHLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAH
SDVCSAQGVRYGPTLKLFKPGQEAVKYQGPRDFQTLENWMLQTLNEEPVTPEPEVEPPSAPE
LKQGLYELSASNFEHLVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCT
QHYELCSGNQVRGYPTLLWFRDGKKVDQYKGKRDLESLEYVESQLQRTETGATETVTPSEA
PVLAEEPEADKGTVLALTENNFDITIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLA
GVKIAEVDCATAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDLHRFVLSQAKDEL
```

Signal sequence:

amino acids 1-32

FIGURE 36

CTTTCTGAGGAACCACAGCAATGAATGGCTTGCATCCTGCTTCGAAGAAACCAATTAT
CCTCCTGGTACTATTCCTTGCAAATTCAGAGTCTGGGCTGGATATTGATAGCCGTCTA
CCGCTGAAGTCTGTGCCACACACACAATTCACCAAGGAGATGATGGTGAAAAAA
GGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGGGACGCATGGGCCGAAAGGAATTAA
AGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCATTGGAAGA
AGGGTACAAAGGGAAAAAGGTTGCTTGAATACCTGGAGAAAAGGCAAAGCAGGTACT
GTCTGTGATTGTGGAAAGAACCGGAAATTGTTGGACACTGGATATTAGTATTGCTCGGCT
CAAGACATCTATGAAGTTGTCAAGAATGTGATAGCAGGATTAGGAACTGAAGAGAAAT
TCTACTACATCGTGCAGGAAGAGAACTACAGGAATCCCTAACCACTGCAGGATTCGG
GGTGGAATGCTAGCCATGCCAAGGATGAAGCTGCCAACACACTCATCGCTGACTATGGTGC
CAAGAGTGGCTTCTTCGGGTGTTCAATTGGCGTGAATGACCTTGAAAGGGAGGGACAGTA
TGTCCACAGACAAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGAACCCAGCGAC
CCCTATGGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTG
CCATCTTACCATGTACTTTGTCATTACAGTTATTGTTATCCATCCTTTTCCTG
ACGTATTTGCTATTTCTGTGACCGTCATTACAGTTATTGTTATCCATCCTTTTCCTG
ATTGTACTACATTTGATCTGAGTCACATAGCTAGAAAATGCTAAACTGAGGTATGGAGCCT
CCATCATCAAAAAAAAAAAAAAA

FIGURE 37

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPGEEG
KHGKVGRMGPKGIKGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPEGKGAGTVCDGRY
RKFVGQLDISIARLKTSMKFVKNVIAGIRETEEKFYIIVQEEKNYRESLTHCIRGGMLAMP
KDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDNTPLQNYSNWNEGEPSDPYGHEDC
VEMLSSGRWNDTECHLTMYFVCEFIKKKK
```

Signal peptide:

amino acids 1-25

FIGURE 38

GGTTCTATCGATTGAACTGGGCCACACTGGCCGGATCCTCTAGAGATCCCTCGACCTCGAC
CCACCGTCCGCTCTCCGCCGTGGAGTGGTGGGGCTGGGTGGGAATGGCGTGT
GCCAGCGACCGCGCTCCCTGGAAGGAGAAGTCTCAGCTAGAACGAGCGGCCCTAGGTTTT
CGGAAGGGAGGATCAGGGATGTTGCGAGCGGCTGGAACCAGACGGTGCCTAGAGGAAGC
GGGCTCCATGGCTGCCCTCTGCTGCTGCCCTGCTGCTGCTGCTACCGCTGCTGCTGCTGA
AGCTACACCTCTGGCCGCAGTTCGCTGGCTTCCGGCGACTTGGCCTTGCGGTGCGAGCT
CTGTGCTGCAAAGGGCTTCTCGAGCTCGGCCCTGGCGGGCTGCCGCCACCCGGAAAGG
TCCCAGGGGGCTGAGCCTGGCCTGGCGCTCGCGAACCTGGCCAGCAGCGGCCGCG
ACACCTTCTCATTACGGCTCGCGCTTAGCTACTCAGAGGCGAGCGAGAGATAAC
AGGGCTGCACGCGCTTACGTGCGTAGGCTGGACTGGGACCCGACGGCGACAG
CGGCAGGGAGCGCTGGAGAAGGCGAGCGGGCAGCGCCGGAGCCGGAGATGAGCAGCG
GAAGCGGCGCGAGTTGCCGGAGGGACGGTGCCAGAGGGTGGAGGAGCCGCCGCCCCT
CTGTACACTGGAGCAACTGTGGCGCTGCTCCTCCCCGCTGGCCAGAGTTCTGTGGCTCTG
GTTGGCTGGCCAAGGCCGCCGCTGCGACTGCCCTGTGCCACGCCCTGCCGGGGCC
CCCTGCTGCACTGCCCTCGCAGCTGCCGCCGCGCTGGTGCTGGGCCAGAGTTCTG
GAGTCCTGGAGCCGACCTGCCCTGAGAGCCATGGGCTCCACCTGTGGCTGCGAGG
CCCAGGAACCAACCTGCTGGAATTAGCGATTGCTGGCTGAAGTGTCCGCTGAAGTGGATG
GCCAGTGCCAGGATAACCTCTTCCCCCAGAGCATAACAGACACGTGCCTGTACATCTC
ACCTCTGGCACCAAGGGCTCCCCAAGGCTGCTGGATCAGTCATCTGAAGATCCTGCAATG
CCAGGGCTTCTATCAGCTGTGGTGTCCACCAGGAAGATGTGATCTACCTGCCCTCCCAC
TCTACCACATGTCCGGTCCCTGCTGGCATCGTGGCTGATGGCATTGGGCCACAGTG
GTGCTGAAATCCAAGTTCTGGCTGGTCAAGTCTGGGAAGATTGCCAGCAGCACAGGGTAC
GGTGTCCAGTACATTGGGAGCTGTGCCGATACCTGTCAACCAGCCCCGAGCAAGGCAG
AACGTGGCATAAGGTCCGGCTGGCAGTGGCAGCGGGCTGCGCCAGATACTGGAGCGT
TTTGTGGCGCTTCGGCCCTGAGGTGCTGGAGACATATGGACTGACAGAGGGCAACGT
GCCACCATCAACTACACAGGACAGCGGGCGCTGTGGGGCTGCTTCTGGCTTACAAGC
ATATCTCCCTCTCCTGATTGCTATGATGTCAACCAGGAGAGCCAATCGGGACCCC
CAGGGCACTGTATGGCCACATCTCAGGTGAGGCCAGGGCTGCTGGGCCCGTAAGCCA
GCAGTCCCCATTCTGGCTATGCTGGCGGCCAGAGCTGGCCCAGGGGAAGTTGCTAAAGG
ATGTCTCCGGCTGGGATGTTTCTTCACACTGGGACCTGCTGGTCTGGATGACCAA
GGTTTCTCCGCTTCCATGATCGTACTGGAGACACCTCAGGTGGAAGGGGAGAATGTGGC
CACAAACGAGGTGGCAGAGGTCTCGAGGCCCTAGATTTCTTCAGGAGGTGAACGTCTATG
GAGTCAGTGTGCCAGGGCATGAAGGCAGGGCTGGAAATGGCAGGCCCTAGTTCTGCTGCTCCCCC
CACGCTTGGACCTTATGCACTACACCCACGTGCTGAGAACCTGCCACCTTATGCCCG
GCCCGATTCTCAGGCTCCAGGAGTCTTGGCCACACAGAGACCTCAAACAGCAGAAAG
TTCGGATGGCAAATGAGGGCTTCGACCCAGCACCCCTGCTGACCCACTGTACGTTCTGGAC
CAGGCTGTAGGTGCCTACCTGCCCTCACAACTGCCGGTACAGCGCCCTCTGGCAGGAAA
CCTTCGAATCTGAGAACTTCCACACCTGAGGCACCTGAGAGAGGAACCTGTGGGTGGGG
CCGTTGCAGGTGACTGGCTGTCAGGGATCTTCTACAGAACCTGCGGTCACTATT
GTAATAAAATGTGGCTGGAGCTGATCCAGCTGTCTGACCTAAAAAAA
AAAAAAAAAGGGCGGCCGACTCTAGAGTCGACCTGCACTAGGGATAACAGGGTAATAAGC
TTGGCCGCCATGGCCAACCTGTTATTGCAG

FIGURE 39

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50913
><subunit 1 of 1, 730 aa, 1 stop
><MW: 78644, pI: 7.65, NX(S/T): 2
MGVCQRTRAPWKEKSQLEAALGFRKGSGMFASGWNQTVPIEEAGSMAALLLPLLLLLPL
LLLKLHLWPQLRWLPADLAFAVRALCCKRALRALARALAAAADPEGPEGGCSLAWRLAELAQO
RAAHTFLIHGSRRFSYSEAERESNRAARAFLRALGWDWGPDGGSGEGSAGEGERAAPGAGD
AAAGSGAEFAGGDGAARGGGAAAPLSPGATVALLPAGPEFLWLWFGTLAKAGLRTAFVPTAL
RRGPLLHCLRSCGARALVLAPEFLESLEPDLPALRAMGLHLWAAGPGTHPAGISDLLAEVSA
EVDPVPGYLSSPQSITDTCLYIFTSGTTGLPKAARIHLKILQCQGFYQLCGVHQEDVIYL
ALPLYHMSGSLLGIVGCMGIGATVVLKSFKSAGQFWEDCQOHRTVVFQYIGELCRYLVNQPP
SKAERGHKVRLAVGSGLRPDTWERFVRRFGPLQVLETYGLTEGNVATINYTGQRGAVGRASW
LYKHIFPFSLIRYDVTGEPIRDPQGHCMATSPGEPLLAVPVSQQSPFLGYAGGPELAQGK
LLKDVFVPGDVFFNTGDLLVCDDQGFLRFHDRTGDTFRWKGENVATTEVAEVFEALDFLQEV
NVYGVTVPGHEGRAGMAALVLRPPHALDLMQLYTHVSENLPYARPRFLRLQESLATTETFK
QQKVRMANEGFDPSTLSDPLVLDQAVGAYLPLTTARYSALLAGNLRI

Type II transmembrane domain:

amino acids 45-65

Other transmembrane domain:

amino acids 379-398

cAMP- and cGMP-dependent protein kinase phosphorylation site
starting at amino acid 136

CUB domain protein motif

amino acids 254-261

putative AMP-binding domain signature

amino acids 332-343

N-glycosylation sites

amino acids 37-40 and 483-486

FIGURE 40

CCTGTGTTAAGCTGAGGTTCCCTAGATCTGTATATCCCCAACACACATACTCCACGCACA
CACATCCCCAAGAACCTCGAGCTCACACCAACAGACACACCGCGCGCATACACACTCGCTCTC
GCTTGTCCATCTCCCTCCGGGGGAGCCGGCGCGCTCCACCTTGCCGCACACTCCGGC
GAGCCGAGCCCGCAGCGCTCCAGGATTCTGGGCTCGGAACCTGGATTGCAGCTCTGAACCC
CCATGGTGGTTTTAAACACTTCTTCTCTTCTCGTTGATTGCACCGTTCCA
TCTGGGGCTAGAGGAGCAAGGCAGCAGCCTCCAGCCAGCCCTGTTGGCTGCCATCGT
CCATCTGGCTTATAAAAGTTGCTGAGCGCAGTCCAGAGGGCTGCCTGCTCGTCCCCTCGG
CTGGCAGAAGGGGTGACGCTGGCAGCGAGGAGCGCCGCTGCCCTGGCGGGCTTT
CGGCTTGAGGGCAAGGTGAAGAGCGCACCGCCGTGGGTTACCGAGCTGGATTGTATG
TTGCACCATGCCTTCTGGATCGGGCTGTGATTCTCCCTCTGGGGCTGCTGCTCTCCC
TCCCCGCCGGGGCGGATGTGAAGGCTGGAGCTGGAGAGGTCCGCCAGCGTACGGTGCC
AAGGGATTGCCTGGGGACATCCCCTACCAAGGAGATGCAGGGAACACTTAAGAATCTG
TCCTCAGGAATATACATGCTGACCAACAGAAATGGAAGACAAGTTAAGCAACAAAGCAAAC
TCGAATTGAAAACCTTGTGGAAGAGACAAGCCATTGTCGCAACACTTTGTGTCAGG
CATAGAAATTGACGAATTTCGAGAGCTCTGGAGAATGCAAGAAAGTCACTAAATGA
TATGTTGACGGACCTATGGCATGCTGTACATGCAAGATTCAAGAGCTTCCAGGACCTCT
TCACAGAGCTGAAAAGGTACTACACTGGGGTAATGTGAATCTGGAGGAATGCTCAATGAC
TTTGGGCTCGGCTCCTGGAACGGATGTTCAGCTGATAAACCTCAGTATCACTCAGTGA
AGACTACCTGGAATGTTGAGCAAATACACTGACAGCTCAAGCCATTGGAGACGTGCC
GGAAACTGAAGATTCAAGGTTACCGCGCCTTCATTGCTGCCAGGACCTTGTCAGGGCTG
ACTGTGGGAGAGAAGTTGCAAACCGAGTTCCAAGGTCAAGCCAAACCCAGGGTGTATCG
TGCCTCATGAAGATGCTGTACTGCCACTGTGGGGCTTCCACTGTGAGGCCCTGCA
ACAACACTGTCTAACGTATGAAGGGCTGTTGGCAAATCAGGCTGACCTGACACAGAG
TGGAATCTGTTATAGATGCAATGCTCTGGTGGCAGAGCGACTGGAGGGCCATTCAACAT
TGAGTCGGTATGGACCCGATAGATGTCAAGATTCTGAAGCATTATGAACATGCAAGAAA
ACAGCATGCAGGTGTCGCAAAGGTCTTCAGGGATGTTGTCAGCCAAACCTGCTCCAGCC
CTCAGATCTGCCGCTCAGCTCTGAAAATTAAATACACGTTCAAGGCTTACAATCCTGA
GGAAAGACCAACAACGCTGTCAGGCACAAGCTGGACCGGCTGGTCACAGACATAAAAGAGA
AATTGAAGCTCTCTAAAAGGTCTGGTCAGCATTACCTACACTATGCAAGGACGAGAGC
GTGACAGCGGGCACGTCACAGGAGGAATGCTGGAACGGCACAGCAAAGCCAGATACTT
GCCTGAGATCATGAATGATGGCTACCAACCAGATCAACAATCCGAGGTGGATGTGGACA
TCACTCGGCCTGACACTTATCAGACAGCAGATTATGGCTCTCCGTGTGATGACCAACAAA
CTAAAAAACGCCCTACAATGGCAATGATGTCATTCCAGGACACAAGTGTGAATCCAGTGG
CTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTCCTCACGGAGTTGAGTTGTCACCA
CAGAGGCCCGCAGTGGATCCCGACCGGGAGAGAGGTGGACTCTCTGCAGCCCAGCGTGGC
CACTCCCTGCTCTCTGGTCTCCTGCAATTGTCCTGGCACTGCAGAGACTGTGCAGATA
ATCTGGGTTTGGTCAGATGAAACTGCATTAGCTATCTGAATGGCCAACACTCACTTCTT
TTCTTACACTCTGGACAATGGACCATGCCACAAAAACTTACGTTTCTATGAGAAGAGAG
CAGTAATGCAATCTGCCTCCCTTTGTTCCCAAAGAGTACCGGGTGCCAGACTGAACAG
CTTCCCTTTCTCAGCTATGTTGGGACCTGTTATTCTAGAGAGAAATTCTTACTCAA
ATTTTCGTACCAGGAGATTCTTACCTCATTGCTTTATGCTGCAGAAGTAAAGGAAT
CTCACGTTGTGAGGGTTTTCTCATTAAAAT

FIGURE 41

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQ
EYTCCTTEMEDKLSQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAEKSNDMF
VRTYGMLYMQNSEVFQDLFTELKRYYTGGNVNLEEMLNDFWARLLERMFQLINPQYHFSEDY
LECVSKYTDQLKPFGDVPRKLKIQVTRAFFIAARTFVQGLTVGREVANRVSKVSPTPGCIRAL
MKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFIDAMLLVAERLEGPFNIES
VMDPIDVKISEAIMNMQENSMQVSAKVFQGCGQPKPAPALRSARSAPENFNTRFRPYNPEER
PTTAAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNNGHSKARYLPE
IMNDGLTNQINNPEVDVDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSG
SGSGCMDDVCPTEFVTEAPAVDPDRREVDSSAAQRGHSLLSWSLTCIVLALQRLCR
```

Signal peptide:

amino acids 1-23

FIGURE 42A

CGGACCGCGTGGCCGGACCGCGTGGCAAAAGAACTCGGAGTCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG
CAGTTGCAGGCCCTGCGCCGGGTGCGCCAACACTACGCAAAGACCAAGCGGGCTCCGCGCGACCGGCCGCGGGGG
TAGGGACCCGGCTTGGCCTCAGGCTCCCTAGCAGCGGGAAAAGGAATTGCTGCCGGAGTTCTGCGGAGGT
GGAGGGAGATCAGGAAACGGCTTCTCCTCACTTCGCCCTGGTGAGTGTGCGGGAGATTGCCAAACGCCTAGG
AAAGGACTGGGGAAAATAGCCCTGGAAAGTGGAGAAGGTGATCAGGAGGCCGCTTCAACTACGGCAGTTATCTG
TCTGATCAGAGCCAGACGCGACCGTCCACTTCGAGTTCTTCCAGGTGTGGGGACCGCAGGACAGACGCCGA
TCCCGCCGCCCTCCGTACCAGCACTCCAGGAGAGTCAGCCTCGCTCCCCAACGTCGAGGGCGCTCTGCCACGA
AAAGTCCCTGTCCACTGTGATTCTCAATTCCCTGCTGGTTTTCTCCAGAGAACTTTGGGTGGAGATATTA
ACTTTTTCTTTTTCTGGTGAAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAAGTGAATGTGC
TGGAGAAGAGCGAGCCCTCCCTGGTCTCCGGAGTCCCATTAAAGCCATCACTCTGGAAGATTAAAGTTGT
CGGACATGGTGACAGCTGAGAGGAGGAGGATTCTGCCAGGTGGAGAGTCTCACCGTCTGGTGCATG
TGTGCGCCCGCAGCGCGCGGGCGCGTGGTCTCCGCGTGAGTCTCACCTGGACCTGAGTGAATGGCTCCA
GGGGCTGTGCGGGGATCCGCTCCGCTTCTCCACAGGCCTGTGCTGCTGGAAAGATGCTAGCAATGGGG
CGCTGGCAGGATTCTGGATCCCTGCCTCCTCACTTATGGTACCTGTCTGGGCCAGGCCTAGAAGAGGAGG
AAGAAGGGCCTACTAGCTCAAGCTGGAGAGAAACTAGAGCCCAGCACAACCTCCACCTCCAGCCCCATCTCA
TTTCATCCTAGCGGATGATCAGGGATTAGAGATGTGGTTACCACGGATCTGAGATTAAACACCTACTCTG
ACAAGCTCGTCCGAAGGAGTTAAACTGGAGAACTACTATGTCCAGCCTATTGACACCCATCAGGAGTCAGT
TTATTACTGGAAAGTATCAGATAACACACCGACTTCAACATTCTATCATAAGACCTACCCACCCAACTGTTAC
CTCTGGACAATGCCACCCCTACCTCAGAAACTGAAGGGAGTGGATATTCAACGCATATGGTCGAAAATGGCACT
TGGGTTTAACAGAAAAGAATGCATGCCACCAGAGAGGATTGATACTTTGGTCCCTTTGGGAAGTG
GGGATTACTATACACACTACAAATGTGACAGTCCTGGGATGTGGCTATGACTTGTATGAAAACGACAATGCTG
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTAGCTCCATA
ACCCCCACAAAGCTATATTCTACTGCCTATCAAGCTGTCATTCACTGCAAGCTCCTGGCAGGTATT
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTCTGCTTAGATGAAGCAA
TCAACAACTGACATTGGCTCTAAAGACTTATGGTTCTATAACACAGCATTATCATTACTCTCAGATAATG
GTGGCCAGCCTACGGCAGGAGGGAGTAAGTGGCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGGAGGATCC
GGGCTGTAGGCTTGTGCATAGCCACTTCTGAAAAAACAGGAAACAGTGTGTAAGGAACCTGTGCACATCACTG
ACTGGTACCCACTCTCATTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT
GGGAGACCATAAGTGAGGGCTTCGCTACCCCGAGTAGATATTGATAACATTGACCCCTATAACACCAAGG
AAAAAAATGGCTCTGGCAGCAGGCTATGGATCTGAAACACTGCAATCAGTCAGCCATCAGAGTGCAGCACTG
GAAATTGCTTACAGGAAATCTGGCTACAGCAGTGGCTCCCTCAGTCTTCAAGCAACCTGGGACCGAACCG
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTCAACATCACAGCCACCCATATGA
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCTACGGAGGCTCTCACAGTTCAACAAAAC
TGCAGTGCCTGGTCAGGTATCCCCCCTAACAGCCCCAGAAGTAACCTAGGCTCAATGGAGGGCTGGGACCATG
GTATAAAGAGGAAACCAAGAAAAAGAAGCCAAGCAAAATCAGGCTGAGAAAAAGCAAAGAAAAGCAAACAAAAAA
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACAGCAGAAATTGGCTCGATAATATCGCTGCCATAAGCGTCA
GGCTTGTGTTCTGCTGTGCACTCCAGAGACTTCTGCCACCTGGCGCCACACTGAAAACGTGCTCTGCTCAGTG
CCAAGGTGCTACTCTGCAAGCCACACTTAGAGAGAGTGGAGATGTTATTCTCTCGCTCCTTGTAGAAAACGTG
GTGAGTCCTGAGTTCCACTGCTGTGCTTCAGTCAGTCAACTGACCAAACACTGCTTGAATTATAGGAGGAGAACAAATA
ACCTACCATCCGCAAGCATGCTAATTGATGGAAGTTACAGGGTAGCATGATTAAACTACCTTGTATAATTAC

FIGURE 42B

AGTCAAAGATTGTGTCACCTCAAAGGCCCTGAAGAATATATTTCTGGTGAATTTGTATGTCTGTCATATGA
CACTGGGTTTTAATTAAATTCTATTATATATAAATATGTTCTTCCTGTGAAAAGCTGTTTCT
CACATGTGAACAGCTGACCTCATTACCATGCGTGAGGAAATGCAAATAAGAATGTTGAGCACACTGCC
ACAATGAATGTAACATTTCTAACACTTACTAGAAGAACATTCACTGATAAAAAACCTAATTTATTTACA
GAAAAATATTTGTTGTTTATAAAAGTTATGCAAATGACTTTATTTTATTCCTGCATACCATTAGAAGA
ATTTATTCATTCTCAAATTATCAAGCACTGTAACTATAAATTAAATGTAACTGTGTAATTCAAGACTA
TAAAAAACATCATTCAAGAAAACCTTATAATCGTCATTGTCATCAAGATTGAAATGTAATAAGATGAATATAT
ATTACTGGAAATTCAATGTTGTGCAGAGTTGAGACAACCTTATTGTTCTATCATAAAACTATTTATGTATCTT
AATTATTAAAATGATTACTTTATGGCACTAGAAAATTACTGTGGCTTTCTGATCTAACTCTAGCTAAAATT
GTATCATTGGCCTAAAAAATTTACTAATAGGAATTGAAGGAATGGTTGCTAACACCACAGTAA
TATAATATGATTTCACAGATAGATGCTCCCTGGCTATGACATGGAGAAAGATTTCCATAATAACTAA
TATTTATATTAGGGTGGTGCACAAACTAGTTGCGGTTTCCATTAAAGTAATAACCTTACTCTTACAAAGT
GGACACTGTGGGAGATAACAGAGAAATGGAAGATAACGGATCCTGCCTGGAGTAGGTAACCTTGCTGGAAACCC
ACATGCAAACGTCAAGGAGAATTAAAGGAGTATTATCAGTAATGAAGTTATCATGGGTCAATGAGCATA
GATTGGTGTGGATCCTGTAGACCCCTGGTTTCTTGAAGTGCCTCTCCTAATGCAGAGGCCCTGAAGCTTAC
AGTATACACTGAAAAGTCACAGATAGCTAGAATTATGATCTTGAAGTTATAACTGTGATCTGAAAATGTGT
GGTGGTATGACAGCATACCATTAAATACATTACACAGCTCAAAGGACTGTGATATAATCCATTATATCAC
AACTCAAAGGACTGTGATATAATCCATTATACACAGCTCACAGTTCTGAAAATGTATAAAGAATCTATAAT
CTAGTACTGAAAATTACTAAATTGGGTAAGATGATTTAAATGATTTAATTAAACATTATTTCTAGAATATAT
GGCTCCATTATTTATAGTGTAAAGTTGTATTCTAAAGTTGTGTTGTCGACAGTATCTTTAAATGAG
TCTTAAAAAATAAGGCATATTGTCATGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAA

FIGURE 43

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAHGPPPSQACVCPGKMLAMGALAGFWILCLLTYGYLSWGQALEEEEGLAQA
GEKLEPSTTSTSQPHLIFILADDQGFRDVGYHGSEIKPTLDKLAEGVKLENYYVQPICTP
SRSQFITGKYQIHTGLQHSIIRPTQPNCLPLDNATLPQKLKEVGYSTHMVGKWHLGFRKEC
MPTRRGFDFFGSLLGSGDYYTHYKCDSPGMCGYDLYENDNAAWDYDNGIYSTQMYTQRVQQ
ILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRYAAMLSCLDEAINVTLA
LKTYGFYNNSSIIYSSDNGGQPTAGGSNWPLRGSKGTWEGGIRAVGFVHSPLLKNKGTVCK
ELVHITDWYPTLISLAEQIIDEDIQLDGYDIWETISEGLRSPRVDILHNIDPYTPRQKMAPG
QQAMGSGTLQSSQPSECSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFST
TSQPTHMRGWTYLTGIQES
```

Important Features:

Signal Peptide:

amino acids 1-37

Sulfatases signature 1.

amino acids 120-132

Sulfatases signature 2.

amino acids 168-177

Tyrosine kinase phosphorylation site.

amino acids 163-169

N-glycosylation sites.

amino acids 157-160, 306-309 and 318-321

FIGURE 44

CGGACCGCGTGGGTGCCAGTGGAGCGGAGGACCCAGCGGCTGAGGAGAGAGGGAGGCCGGC
TTAGCTGCTACGGGGTCCGGCGCCCTCCCGAGGGGGCTCAGGAGGAGGAAGGAGGAC
CCGTGCGAGAATGCCTCTGCCCTGGAGCCTGCGCTCCGCTGCTCTCCTGGTGGCAG
GTGGTTTGGGAACGCGGCCAGTGCAGGCATCACGGTTGTTAGCATCGGCACGTCAGCCT
GGGGTCTGTCACTATGGAACAACTGGCCTGCTGCTACGGCTGGAGAAGAACAGCAAGGG
AGTCTGTGAAGCTACATGCGAACCTGGATGTAAGTTGGTGAGTGCCTGGGACCAAACAAAT
GCAGATGCTTCCAGGATACACCGGGAAAACCTGCAGTCAGATGTGAATGAGTGTGGAATG
AAACCCCGGCATGCCAACACAGATGTGAATACACACCGAAGCTACAAGTGCCTTGCCT
CAGTGGCCACATGCTCATGCCAGATGCTACGTGTGAACCTAGGACATGTGCCATGATAA
ACTGTCAGTACAGCTGTGAAGACACAGAAGAAGGGCCACAGTGCCTGTGTCATCCTCAGGA
CTCCGCCTGGCCCCAAATGGAAGAGACTGTCTAGATATTGATGAATGTGCCTCTGGTAAAGT
CATCTGTCCCTACAATCGAAGATGTGAACACATTGGAAGCTACTACTGCAAATGTCACA
TTGGTTTGAAC TGCAATATATCAGTGGACGATATGACTGTATAGATATAAATGAATGTACT
ATGGATAGCCATACGTGCAGCCACCATGCCATTGCTCAATACCCAAAGGGCTTCAAGTG
TAAATGCAAGCAGGGATATAAAGGCAATGGACTTCGGTGTCTGCTATCCCTGAAAATTCTG
TGAAGGAAGTCCCTCAGAGCACCTGGTACCATCAAAGACAGAATCAAGAAGTTGCTTGCAC
AAAAACAGCATGAAAAAGGCAAAATTAAAAATGTTACCCAGAACCCACCAGGACTCC
TACCCCTAAGGTGAACTTGCAGCCCTCAACTATGAAGAGATAGTTCCAGAGGGCGGGAACT
CTCATGGAGGTAAAAAGGAATGAAGAGAAATGAAGAGGGCTTGAGGATGAGAAAAGAG
AAGAGAAAGCCTGAAGAATGACATAGAGGAGCGAACGCTGCGAGGGAGATGTGTTTCCCT
AAGGTGAATGAAGCAGGTGAATTGGCCTGATTGGTCAAAGGAAAGCGCTAACTTCAA
ACTGGAACATAAAGATTTAAATATCTGGTTGACTGCAGCTCAATCATGGATCTGTGACT
GGAAACAGGATAGAGAAGATGATTTGACTGGAATCCTGCTGATCGAGATAATGCTATTGGC
TTCTATATGGCAGTTCCGGCCTGGCAGGTACAAGAAAGACATTGGCGATTGAAACTTCT
CCTACCTGACCTGCAACCCAAAGCAACTCTGTTGCTCTTGATTACCGCTGGCCGGAG
ACAAAGTCGGAAACTCGAGTGGTGTGAAAAACAGTAACAAATGCCCTGGCATGGAGAAG
ACCACGAGTGAGGATGAAAAGTGGAAAGACAGGGAAAATTCAAGTGTATCAAGGAACGTG
TACCAAAAGCATCATTGGTAAAGCAGAACGTGGCAAGGGAAAACCGCGAAATCGCAGTGG
ATGGCGTCTTGTGTTGAGGCTTATGTCCAGATAGCCTTTATCTGTGGATGACTGAATG
TTACTATCTTATATTGACTTGTATGTCAGTTCCCTGGTTTTTGATATTGCATCATAG
GACCTCTGGCATTAGAATTACTAGCTGAAAAATTGTAATGTACCAACAGAAATATTATTG
TAAGATGCCTTCTTGATAAGATATGCCAATATTGCTTAAATATCATATCACTGTATCT
TCTCAGTCATTCTGAATCTTCCNCATTATAATTATAAAATNTGAAANGTCAGTTATCTC
CCCTCCTCNGTATATCTGATTTGTATANGTANGTTGATGNCTCTCTACAACATTTCTA
GAAAATAGAAAAAAAGCACAGAGAAATGTTAACTGTTGACTCTTATGATACTTCTTGG
AACTATGACATCAAAGATAGACTTTGCCTAAGTGGCTTAGCTGGTCTTCATAGCCAAC
TTGTATTTAATTCTTGTAAATAATAA

40012002-102407

FIGURE 45

MPLPWSLALPLLLSWVAGGFGNAASARHHGLLASARQPGVCHYGTKLACCYGWRRNSKGVCE
ATCEPGCKFGEVGPNKCRCFPGYTGKTCSDVNECGMKPRPCQHRCVNTHGSYKCFCLSGH
MLMPDATCVNSRTCAMINCQYSCEDTEEGPQCLCPSSGLRLAPNGRDCLDIDEASGKVICP
YNRRCVNTFGSYYCKCHIGFELQYISGRYDCIDINECTMDSHTCSHHANCFTQGSFKCKCK
QGYKGNGLRCSAIPENSVKEVLRAPGTIKDRIKLLAHKNSMKKAKIKNVTPETRPTPK
VNLQPFNYEEIVSRGGNSHGGKGNEEK

Signal peptide:

amino acids 1-21

EGF-like domain cysteine pattern signature.

amino acids 80-91

Calcium-binding EGF-like domains

amino acids 103-124, 230-251 and 185-206

FIGURE 46

GGGAGCTGCTGCTGTCGGCTGCTGGCTGTGCGCGCTGCTCCTGCTCTGGTGCAGCTGCTG
CGCTTCCTGAGGGCTGACGGCGACCTGACGCTACTATGGGCCGAGTGGCAGGGACGACGCC
AGAATGGGAGCTGACTGATTGGGTGGTGTGGGTGACTGGAGCCTCGAGTGGAAATTGGTGAGG
AGCTGGCTTACCAAGTTGTCTAAACTAGGAGTTCTCTTGCTGTCAGCCAGAAGAGTCAT
GAGCTGGAAAGGGTGAAAAGAAGATGCCTAGAGAATGGCAATTAAAAGAAAAAGATATACT
TGTTTGCCCCCTGACCTGACCGACACTGGTCCCAGTGAAGCGGCTACCAAAGCTGTTCTCC
AGGAGTTGGTAGAATCGACATTCTGGTCAACAATGGTCCAATGTCCCAGCGTTCTGTGC
ATGGATACCAGCTTGGATGTCTACAGAAAGCTAACAGACTACTTAGGGACGGTGTC
CTTGACAAAATGTGTTCTGCCTCACATGATCGAGAGGAAGCAAGGAAAGATTGTTACTGTGA
ATAGCATCCTGGGTATCATATCTGTACCTCTTCCATTGGATACTGTGCTAGCAAGCATGCT
CTCCGGGGTTTTTAATGCCCTCGAACAGAACCTGCCACATACCCAGGTATAATAGTTTC
TAACATTGCCAGGACCTGTCAATCAAATATTGGAGAATTCCCTAGCTGGAGAAGTCA
CAAAGACTATAGGCAATAATGGAGACCAGTCCCACAAGATGACAACCAGTCGTTGTGCGG
CTGATGTTAACATCAGCATGGCAATGATTGAAAGAAGTTGGATCTCAGAACAAACCTTCTT
GTTAGTAACATATTGGCAATACATGCCAACCTGGCCTGGTGGATAACCAACAAGATGG
GGAAGAAAAGGATTGAGAACTTTAACAGAGTGGTGTGGATGCAGACTCTTCTTATTAAAATC
TTAACAGACAAAACATTGACTGAAAAGAGCACCTGTACTTTCAAGCCACTGGAGGGAGAAATG
GAAAACATGAAAACAGCAATCTCTTATGCTTCTGAATAATCAAAGACTAATTGTGATT
ACTTTTAATAGATATGACTTGTCTCAACATGGAATGAAATAAAAATAATAATAAAAAG
ATTGCCATGAATCTTGCAGAA

20140324-10420409

FIGURE 47

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA36343
><subunit 1 of 1, 289 aa, 1 stop
><MW: 32268, pI: 9.21, NX(S/T): 0
M V V W V T G A S S G I G E E L A Y Q L S K L G V S L V L S A R R V H E L E R V K R R C L E N G N L K E K D I L V L P L D L
T D T G S H E A A T K A V L Q E F G R I D I L V N N G G M S Q R S I C M D T S L D V Y R K L I E L N Y L G T V S L T K C V L
P H M I E R K Q G K I V T V N S I L G I I S V P L S I G Y C A S K H A L R G F F N G L R T E L A T Y P G I I V S N I C P G P
V Q S N I V E N S L A G E V T K T I G N N G D Q S H K M T T S R C V R L M L I S M A N D L K E V W I S E Q P F L L V T Y L W
Q Y M P T W A W W I T N K M G K K R I E N F K S G V D A D S S Y F K I F K T K H D
```

Important Features:

Signal Peptide:

amino acids 1-31

Transmembrane domain:

amino acids 136-157

Tyrosine kinase phosphorylation site.

106-113 and 107-114

Homologous region to Short-chain alcohol dehydrogenase

amino acids 80-90, 131-168, 1-13 and 176-185

FIGURE 48

GCAGCTGGGACCGCCATCAGCTTCGCGCTTCTCCTCCAGGTGGGCAGGGTTTC
GGGCTGGTGGAGCATGTGCTGGACAGGACAGCATCCTCAATCCAACAGCATATTGG
TTGCATCTTCTACACACTACAGCTATTGTTAGGTTGCCTGCGGACACGCTGGGCCTCTGTCC
TGATGCTGAGCTCCCTGGTCTCGCTGGTCTGTCTACCTGGCCTGGATCCTGTTC
TTCGTGCTCTATGATTCTGCATTGTTGTATCACACACCTATGCTATCAACGTGAGCCTGAT
GTGGCTCAGTTCCGGAAGGTCCAAGAACCCCAGGGCAAGGCTAACAGGGACTGAGCCTCA
ACCCAAGCCAGGCTGACCTCATCTGCTTGCTTGCTTCAAGCCGCTCAGCGTGCCTGTG
GACAGCGTGGCCCCGGCCCCCAAGCCTCAGGAGGGCAACACAGTCCCTGGCGAGTGGCCC
TGGCAGGCCAGTGTGAGGAGGCAAGGAGCCCACATCTGCAGCGGCTCCCTGGTGGCAGACAC
CTGGGTCCACTGCTGCCACTGCTTGAAAGGCAGCAGCAACAGAACTGAATTCTGGT
CAGTGGTCTGGTTCTCGCAGCGTGGGACTCAGCCCTGGGCCGAAGAGGTGGGGTG
GCTGCCCTGCAGTTGCCAGGGCTATAACCACTACAGCCAGGGCTCAGACCTGGCCCTGCT
GCAGCTGCCACCCACGACCCACACACCCCTCTGCCTGCCAGCCGCATCGCTTCC
CCTTGAGCCTCTGCTGGGCACTGGCTGGGATCAGGACACCAGTGTAGCTCCTGGGACC
CTACGCAATCTGCCCTGCGTCTCATCAGTCGCCCCACATGTAACGTATCTACAACCAGCT
GCACCAAGCGACACCTGCTCAACCCGGCCCTGGGATGCTATGTGGGGCCCCAGCCTG
GGGTGCAGGGCCCTGTCAGGGAGATTCCGGGGCCCTGTGCTGTGCCCTGAGCCTGACGGA
CACTGGGTCAGGCTGGCATCATCAGCTTGATCAAGCTGTGCCAGGAGGACGCTCCTGT
GCTGCTGACCAACACAGCTGCTCACAGTCCCTGGCTGCAGGCTCGAGTTCAGGGGCAGCTT
TCCTGGCCCAGAGCCCAGAGACCCGGAGATGAGTGTAGGAGACAGCTGTGTAGCCTGTGGA
TCCTTGAGGACAGCAGGTCCCCAGGCAGGAGCACCCCTCCCCATGCCCTGGGAGGCCAGGCT
GATGCACCAGGGACAGCTGCCCTGTGGGGAGGCCCTGGTGTAGAGGAGGCGGTGCTAACTG
CTGCCCACTGCTTCATTGGCGCCAGGCCAGAGGAATGGAGCGTAGGGCTGGGACCAAGA
CCGGAGGAGTGGGCCTGAAGCAGCTCATCCTGCATGGAGCCTACACCCACCCCTGAGGGGG
CTACGACATGCCCTCTGCTGGCCAGCCTGTGACACTGGGAGCCAGCCTGCC
TCTGCCCTGCCCTATCCTGACCACCACTGCCCTGATGGGAGCGTGGCTGGGTTCTGGGACGG
GCCCGCCCAGGAGCAGGCATCAGCTCCCTCAGACAGTGCCGTGACCCCTCTGGGCTAG
GGCCTGCAGCCGGCTGCATGCAGCTCTGGGGTGATGGCAGCCCTATTCTGCCGGGATGG
TGTGTACCAAGTGTGGGTGAGCTGCCAGCTGTGAGGGCCTGTCTGGGCACCACTGGT
CATGAGGTGAGGGGCACATGGTCTCTGGCGGGCTGCACAGCTCGAGATGCTTGCCAAGG
CCCCGCCAGGCCGGCGGTCTCACCGCGCTCCCTGCCTATGAGGACTGGGTAGCAGTTGG
ACTGGCAGGTCTACTTCGCCAGGAACCAGAGCCGAGGCTGAGCCTGGAAGCTGCC
AACATAAGCCAACCAACCAGCTGCTGACAGGGACCTGCCATTCTCAGGACAAGAGAATGC
AGGCAGGCAAATGGCATTACTGCCCTGTCCTCCCCACCCCTGTATGTGTGATTCCAGGCAC
CAGGGCAGGCCAGAACGCCAGCAGCTGTGGGAAGGAACCTGCCTGGGCCACAGGTGCCA
CTCCCCACCCCTGCAGGACAGGGGTGTCTGTGGACACTCCCACACCAACTCTGCTACCAAGC
AGGCCTCTCAGCTTCTCTCTTACTCTTCAGATACAATCACGCCAGCCACGTTGTT
TGAAAATTCTTTGGGGGCAGCAGTTCTTTAAACTAAATAATTGTTAC
AAAATAAA

FIGURE 49

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40571
MLLSSLVSLAGSVYLAWLFFVLYDFCIVCITYAINVSLMWSFRKVQEPMQGKAKRHGNTV
PGEWPWQASVRRQGAHICSGSLVADTWVLTAAHCFEKAATELNSWSVVLGSLQREGLSPGA
EEVGVAALQLPRAYNHYSQGSIDLALLQLAHPTTHTPLCLPQPAHRFPFGASCWATGWDQDTS
DAPGTLRNLRLRLISRPTCNCIYNQLHQRHLSNPARPGMLCGGPQPGVQGPGCQGDGGPVLC
LEPDGHWVQAGIISFASSCAQEDAPVLLNTAAHSSWLQARVQGAFLAQSPEPEMSDEDS
CVACGSLRTAGPQAGAPSPWPWEARLMHQQLACGGALVSEEAVLTAAHCFIGRQAPEEWSV
GLGTRPEEGLKQLILHGAYTHPEGGYDMALLLLAQPVTLGASLRPLCLPYPDHHLPDGERG
WVLGRARPGAGISSLQTVPTLLGPRACSLHAAAPGGDGSPILPGMVCTSAVGELPSCEGLS
GAPIVHEVRGTWFLAGLHSFGDACQGPAPAVFTALPAYEDWVSSLDWQVYFAEEPEPEAEP
GSCLANISQPTSC
```

Important features:

Signal peptide:

amino acids 1-15

Homologous region to Serine proteases, trypsin family

amino acids 79-95, 343-359 and 237-247

N-glycosylation sites.

amino acids 37-40 and 564-567

Kringle domains

amino acids 79-96, 343-360 and 235-247

FIGURE 50

CGGGCCGCCCGCCCCCGCCCCATTGGGCGGGCCTCGCTGGCGGGCGACTGAGCCAGGCTGG
GCCGCGTCCCTGAGTCCCAGAGTCGGCGGGCGGGCAGGGCAGCCTCACACGGGAG
CCCAGCTGTCAAGCCGCCTCACAGGAAGATGCTGCGTCGGCGGGCAGCCCTGGCATGGTGT
GCATGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGG
TCCCTGAAGACCCAGTGGTGGCACTGGTGGCACCGATGCCACCCGTGCTGCTCCTCTCC
CCTGAGCCTGGCTTCAGCCTGGCACAGCTAACCTCATCTGGCAGCTGACAGATAACAAACA
GCTGGTGCACAGCTTGCTGAGGGCCAGGACCAGGGCAGCGCTATGCCAACCGCACGGCCC
TCTTCCCGGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGTGCCTGGCG
GACGAGGGCAGCTTCACCTGCTCGTGAGCATCCGGATTTCGGCAGCGTGCCTCACGCT
GCAGGTGGCCGCTCCCTACTCGAAGCCCAGCATGACCCTGGAGCCAAACAAGGACCTGCGGC
CAGGGACACGGTGACCATCACGTGCTCAGTACCAAGGGTACCCGTGAGGCTGAGGTGTT
TGGCAGGATGGCAGGGTGTGCCCTGACTGGCAACGTGACCACGTCGCAGATGGCCAACGA
GCAGGGCTTGTGATGTGCACAGCGTCTGCGGGTGGTGTGGTGCATGGCACCTACA
GCTGCCTGGTGCACACCCCGTGCAGCAGGATGCGCACRGCTCTGTCAACATCACAGGG
CAGCCTATGACATTCCCCCAGAGGCCCTGTGGGTGACCGTGGGCTGTCTGTCTCAT
TGCACTGCTGGTGGCCCTGGCTTCGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGG
AGAATGCAGGAGCTGAGGACCAGGATGGGAGGGAGAAGGCTCCAAGACAGCCGTGAGCCT
CTGAAACACTCTGACAGCAAAGAAGATGATGGACAAGAAATAGCCTGACCATGAGGACCAAG
GAGCTGCTACCCCTCCCTACAGCTCCTACCCCTGGCTGCAATGGGCTGCACTGTGAGCCC
TGCCCCAACAGATGCATCCTGCTCTGACAGGTGGCTCCTCTCAAAGGATGCGATAACAC
AGACCACTGTGCAGCCTTATTCCTCAATGGACATGATTCCAAGTCATCCTGCTGCCTTT
TTCTTATAGACACAATGAACAGACCACCAACCTTAGTTCTCTAAGTCATCCTGCCTGCT
GCCTTATTCACAGTACATACATTTCTTAGGGACACAGTACACTGACCACATCACCACCC
TTCTTCCAGTGCCTGGGACCATCTGGCTGCCTTTCTCAAAGGATGCAATATTCA
CTGACTGACCCCTGCCTTATTCACCAAAGACACGATGCATAGTCACCCGGCCTGTT
TCCAATGGCCGTGATAACTAGTGTATGTTAGCCCTGCTCCACCTGCATAGAATCTT
TCTTCTCAGACAGGGACAGTGCCTCAACATCTGGAGTCTAGAAGCTGTTCT
CCCTCCTCCTCCCTGCCCAAGTGAAGACAGGGCAGGGCAGGAATGCTTGGGACACCG
AGGGGACTGCCCTGCCCAAGTGAAGACAGGGCAGGGCAGGAATGCTTGGGACACCG
TTGCCTCTGGCCAGCTCCTGGCCTCTGGTAGAGTGAGACTTCAGACGTTCTGATGCCTCCG
GATGTCATCTCCTGCCCTGCCCAAGGAATGGAAGATGTGAGGACTTCTAATTAAATGTGGGAC
TCGGAGGGATTTGTAAACTGGGGTATATTTGGGAAAATAATGTCTTGTAAAAAAA
AAAAAAA

FIGURE 51

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHGAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPFGFSLAQ
LNLIWQLTDKQLVHSFAEGQDQGSAYANRTALFPDLLAQGNASLRLQRVRVADEGSFTCFV
SIRDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDTVTITCSSYQGYPEAEVFWQDGQGVPL
TGNVTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEA
LWVTVGLSVCLIALLVALAFVCWRKIKQSCEEENAGAEDQDGEGEGSKTALQPLKHSDSKED
DGQEIA
```

Important features:

Signal peptide:

amino acids 1-28

Transmembrane domain:

amino acids 251-270

N-glycosylation site.

amino acids 91-94, 104-107, 189-192 and 215-218

Homologous region to Immunoglobulins and MHC

amino acids 217-234

FIGURE 52

TTCTGTGACCCCTTGAGAAAAGAGTTGGTGGTAAATGTGCCACGTCTCTAAGAAGGGGGAGTC
CTGAACCTGTCTGAAGCCCTGTCCGTAAGCCTGAACTACGTTCTTAAATCTATGAAGTCG
AGGGACCTTCGCTGTTGTAGGGACTTCTTCCTTGCTTCAGCAACATGAGGCTTTCT
TGTGGAACGCGGTCTGACTCTGTCGTCACTTCTTGATTGGGGCTTGATCCCTGAACCA
GAAGTGAAAATTGAAGTTCTCCAGAAGCCATTCACTGCGCATCGCAAGACCAAAGGAGGGGA
TTTGATGTTGGTCCACTATGAAGGCTACTTAGAAAAGGACGGCTCTTATTCACTCCACTC
ACAAACATAACAATGGTCAGCCATTGGTTACCCCTGGCATCCTGGAGGCTCTCAAAGGT
TGGGACCAAGGGCTTGAAGGAATGTGTAGGAGAGAAGAGAAAGCTCATCATTCCCTGC
TCTGGGCTATGGAAAAGAAGGAAAAGGTTAAAGTACACTGATATTAAATA
TTGATCTCCTGGAGATTGAAATGGACCAAGATCCCAGAATGGATCTTCAAGAAATGGATCTT
AATGATGACTGGAAACTCTCTAAAGATGAGGTTAAAGCATATTAAAGAAGGAGTTGAAAAA
ACATGGTGCAGGTGGTGAATGAAAGTCATCATGATGCTTGGTGGAGGATATTGATAAAG
AAGATGAAGACAAAGATGGTTATATCTGCCAGAGAATTACATATAAACACGATGAGTTA
TAGAGATACATCTACCCCTTTAATATAGCACTCATCTTCAAGAGAGGGCAGTCATCTTAA
AGAACATTTATTTTATACAATGTTCTTCTGCTTGTGTTTATTTTATATTTT
CTGACTCCTATTTAAAGAACCCCTTAGGTTCTAAGTACCCATTCTTCTGATAAGTTATT
GGGAAGAAAAGCTAATTGGTCTTGAATAGAAGACTTCTGGACAATTTCACAG
ATATGAAGCTTGTACTTCTCACCTATAAATTAAATTTGCAACTGGAAATATACC
ACGACATGAGACCAGGTTATAGCACAAATTAGCACCCATTCTGCTTCCCTCATTTTC
TCCAAGTTAGAGGTCAACATTGAAAAGCCTTGTCAATAGCCCAGGCTGCTATTTCAT
GTTATAATGAAATAGTTATGTGTAACTGGCTCTGAGTCTGCTTGAGGACCAAGGAAAA
TGGTTGTTGGACCTGACTTGTAAATGGCTACTGCTTACTAAGGAGATGTGCAATGCTGAAG
TTAGAAACAAGGTTAATAGCCAGGCATGGTGGCTATGCCGTAAATCCCAGCACTTGGGAG
GCTGAGGCGGGCGGATCACCTGAGGTTGGGAGTTGAGACCAGCCTGACCAACACGGAGAAA
CCCTATCTCTACTAAAATACAAAGTAGCCCCGGCGTGGTGTGCGTGCCTGTAATCCCAGCT
ACCCAGGAAGGCTGAGGCGGAGAACATCACTGAAACCGAGGCCAGGTTGCGGTAAAGCCAG
ATCACCTNCAGCCTGGACACTCTGTCGAAAAAAAGAAAAGAACACGGTTAATACCATATNA
ATATGTATGCATTGAGACATGCTACCTAGGACTTAAGCTGATGAAAGCTTGGCTCTAGTGT
TGGTGGCTATTATGATAAAATAGGACAAATCATTATGTGTGAGTTCTTGTAATAAAATG
TATCAATATGTTATAGATGAGGTAGAAAGTTATTTATATTCAATATTTACTTCTTAAGGC
TAGCGGAATATCCTCTGGTTCTTAATGGGTAGTCTATAGTATATTACTACAATAACA
TTGTATCATAAGATAAAAGTAGTAAACCAGTCTACATTCTCCATTCTGTCATCAAAAC
TGAAGTTAGCTGGGTGTGGTGGCTATGCCGTAAATCCCAGCACCTTGGGGCCAAGGAGGG
TGGATCACTTGAGATCAGGAGTCAAGACCCAGCCTGGCCAACATGGTGAACACCTTGTCTCTA
CTAAAAATACAAAATTAGCCAGGCAGTGGTGGTGCACACCTGTAGTCCCAGCTACTCGGGAG
GCTGAGACAGGAGATTGCTTGAACCCGGGAGGGAGGTTGCACTGAGCCAAGATTGTGCC
ACTGCACTCCAGCCTGGTGACAGAGCAAGACTCCATCTCAAAAAAAAAAAGAAGCAGA
CCTACAGCAGCTACTATTGAATAAACCTATCTGGATTTT

FIGURE 53

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44194
><subunit 1 of 1, 211 aa, 1 stop
><MW: 24172, pI: 5.99, NX(S/T): 1
MRLFLWNAVLTLFVTSLIGALIPEPEVKIEVLQKPFICHRKTGGDLMLVHYEGYLEKD GSL
FHSTHKHNNGQPIWFTLGLILEALKGWDQGLKGMCVGEKRKLIIPPALGYGKEGKGKIPPEST
LIFNIDLLEIRNGPRSHESFQEMDLNDDWKL SKDEVKAYLKKEFEKHGAVVNESHDALVED
IFDKEDEDKDGFISAREFTYKHDEL
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 176-179

Casein kinase II phosphorylation site.

amino acids 143-146, 156-159, 178-181 and 200-203

Endoplasmic reticulum targeting sequence.

amino acids 208-211

FKBP-type peptidyl-prolyl cis-trans isomerase

amino acids 78-114 and 118-131

EF-hand calcium-binding domain.

amino acids 191-203, 184-203 and 140-159

S-100/ICaBP type calcium binding domain

amino acids 183-203

FIGURE 54

AATAAAGCTCCTTAATGTTGATATGCTTTGAAGTACATCCGTGCATTTTTTTAGCAT
CCAACCATTCCCTCCCTGTAGTTCTCGCCCCCTCAAATCACCTCTCCGTAGCCCACCGA
CTAACATCTCAGTCTCTGAAA**ATGC**CACAGAGATGCCTGGTACCTCGCCCTGCCTTCAGCCT
CACGGGGCTCAGTCTCTTCTCTTGGTGCACCGAGGACGGAGCATGGAGGTACAGTAC
CTGCCACCTCAACGTCCATGGCTCTGACGCCGCCTGCCCTGCACCTCAACTCTGC
TACACAGTGAACCACAAACAGTTCTCCCTGAACTGGACTTACCAAGGAGTGCAACAACTGCTC
TGAGGAGATGTTCTCCAGTCCGATGAAGATCATTAACCTGAAGCTGGAGCGGTTCAAG
ACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTATGCTGAGAAACGTG
CAGCCGGAGGATGAGGGGATTACAACGTACATCATGAACCCCCCTGACGCCACCGTGG
CCATGGCAAGATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCAGGGACTCCACGGTGG
CCGTGATTGTGGGTGCCTCCGTGGGGCTTCTGGCTGTGGTCATCTTGGTGTGATGGTG
GTCAAGTGTGTGAGGAGAAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACGGAGGA
GGAGGGCAAGACGGACGGTGAAGGCAACCCGGATGATGGCGCCAAG**TAGTGGGTGGCGGCC**
CTGCAGCCTCCCGTGTCCCGTCTCCTCCCTCTCCGCCCTGTACAGTGACCTGCCTGCTCG
CTCTGGTGTGCTCCCGTACCTAGGACCCCCAGGGCCACCTGGGGCTCTGAACCCCCG
ACTTCGTATCTCCACCTGCACCAAGAGTGACCCACTCTTCCATCCGAGAAACCTGCCA
TGCTCTGGGACGTGTGGGCCCTGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCTGG
GGGGAGGCAGGAGGCACATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGCAGGGTAGA
GGAGGGGCCGCTGTCACCTGCCAGTGCTTGGCTCAGAGAGGACCTGGTGG
GGAGGGAGGGCTTCTGTGCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTG
TGCTCCTCCCTGCTCCAGCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTCCCTGA
AACTGGAGGGCATGTTAAAGGGATGACTGTGCATTCCAGGGCACTGACGGAAAGCCAGGG
CTGCAGGCAAAGCTGGACATGTGCCCTGGCCAGGAGGCCATGTTGGGCCCTCGTTCCATT
GCTAGTGGCCTCCTGGGCTCTGGCTCTTAATCCCTAGGACTGTGGATGAGGCCAG
ACTGGAAGAGCAGCTCCAGGTAGGGGCCATGTTCCAGCGGGGACCCACCAACAGAGGCC
AGTTCAAAGTCAGCTGAGGGCTGAGGGTGGGCTCCATGGTGAATGCAGGTTGCTGCAG
GCTCTGCCCTCTCCATGGGTAACCACCCCTGCCCTGGCAGGGCAGCCAAGGCTGGAAAT
GAGGAGGCCATGCACAGGGTGGGCAGCTTCTGGCTACAGAGAGGAAAGGCTGAGGCCAG
GCCCTTGGTGGGTTCCACCTGGCTTGGCTACAGAGAGGAAAGGCTGAGGCCAG
GCATAAGGGGAGGCCTGGAACCTGAGCTGCCATGCCAGCCCTGTCCATCTGCCACG
CTACTCGCTCCTCTCCCAACAACTCCCTCGTGGGACAAAAGTACAATTGTAGGCCAGGC
ACAGTGGCTACGCCGTGAATCCCAGCATTGGAGGCCAAGGCGGGTGGATTACCTCCAT
CTGTTAGTAAAGGGCAAACCCATCTCTACTAAAATACAAGAATTAGCTGGCGTG
GTGGCGTGTGCGCTGTAATCCCAGCTATTGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCG
GGAAGCAGAGGTTGCAGTGAACTGAGATAGTGAATGTGCAATTGCACTGCAATTGCACTGGGTGAC
ATAGAGAGACTCCATCTCAAAAAAA

FIGURE 55

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415
<subunit 1 of 1, 215 aa, 1 stop
<MW: 24326, pI: 6.32, NX(S/T): 4
MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQ
FSLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRNQPEDEGI
YNCYIMNPPDRHRGHGKIHQLQVLMEEPPRDSTVAVIVGASVGGFLAVVILVLMVVKCVRRK
KEQKLSTDDLKTEEEGKTDGEGNPDDGAK
```

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 161-179

Immunoglobulin-like fold:

amino acids 83-127

N-glycosylation sites.

amino acids 42-45, 66-69 and 74-77

FIGURE 56

GTTGTATATGTCCTGAAGTACATCCGTGCATTTTTAGCATCCAACCATCCTCCCTGTA
GTTCTCGCCCCCTCAAATCACCTTCTCCCTAGCCCACCCNACTAACATCTCAGTCTCTGAA
AATGCACAGAGATGCCTGGCTACCTCGCCCTGCCCTCAGCCTCACGGGGCTCAGTCTCTTT
TCTCTTGGTGCCACCAGGACGGAGCATGGAGGTCCACAGTACCTGNCCACCCCTAACGTCC
TCAATGGCTCTGACGCCCGCCTGCCCTGCCCTCAACTCCTGCTACACAGTGAACCACAAAC
AGTTCTCCCTGAACTGGACTTACCAAGGAGTGCAACAACTGCTCTGAGGAGATGTTCTCCAG
TTCCGCATGAAGATCATTAACCTGAAGCTGGAGCGGTTCAAGACCGCGTGGAGTTCTCAGG
GAACCCCAGCAAGTACGATGTGTCGGTATGCTGAGAACGTGCAGCCGGAGGATGAGGGGA
TTTACAACGTACATCATGAACCCCCC

FIGURE 57

TCACGGGGCTCATCTTTCTTTGGTCCCACCAGGACGGACATGGAGGTNCACATA
CCTGCCACCCCTAACGTCTCAATGGCTTGACGCCGCCTGCCCTGCACCTCAACTCCNG
CTACACAGTGAACCACAAACAGTTCTCCCTGAACGGATTACCAAGGAGTGCACAAACTGGC
TCTGAGGAGATGTTCCCTCCAGTTCCGCATGGAAGATCATTAAACCTGAAAGCTGGAAGCGG
TTTCAAGAACCGCGTGGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTCGGTGATGC
TGAGAAACGTGCAGCCGGAGGATGAGGGGATTACAACGTACATCATGAACCCCCC

FIGURE 58

TGGCGCGACCGTCGTACACCATGGGCTCCACCTCCGGCCCTTACCGTGTGGGCTGCTCCGGATGGCTCCCTGT
TCCTCTTGCTGCTGCTAATGCTGCTCGGGACCCAGCGCTCCGGCGACGTCAACCCCCAGTGGTGTGGTCC
CTGGTGATTGGGTAACCAACTGGAAGCCAAGCTGGACAAGCCGACAGTGGTGCACATACCTCTGCTCCAAGAAGA
CCGAAAGCTACTTCACAATCTGGCTGAACCTGGAACGCTGCTGCTGTCAATTGACTGCTGGATTGACAATA
TCAGGCTGGTTACAACAAAACATCCAGGGCCACCCAGTTCTGTATGGTGTGGATGTACGTGTCCCTGGCTTTG
GGAAGACCTTCTCACTGGAGTTCTGGACCCAGCAAAAGCAGCGTGGGTTCTATTTCACACCATGGTGGAGA
GCCTTGTGGGCTGGGCTACACACGGGTGAGGATGTCCGAGGGGCTCCCTATGACTGGCGCCAGGCCAAATG
AAAACGGGCCCTACTTCTGGCCCTCCGCGAGATGATCGAGGAGATGTACCGAGCTGTATGGGGCCCGTGGTGC
TGGTTGCCACAGTATGGGCAACATGTACACGCTCTACTTCTGCAAGCGGAGCCGAGGCCCTGGAAGGACAAGT
ATATCCGGGCTTCGTGTCAGGGTGCCTGGGGGGCGTGGCCAAGACCTGCGCGTCTGGCTTCAGGAG
ACAACAACCGGATCCCAGTCATCGGGCCCTGAAGATCCGGGAGCAGCAGCGGTCAAGCTGCTCCACAGCTGGC
TGCTGCCCTACAACACTACACATGGTCACCTGAGAAGGTGTTCTGCAAGACACCCACAATCAACTACACACTGCGG
ACTACCGCAAGTTCTCCAGGACATCGGCTTGAGAAGATGGCTGGCTCATGCGGAGGACACAGAAGGGCTGGTGG
AAGCCACGATGCCACCTGGCGTGCAGCTGCACTGCCTCTATGGTACTGGCGTCCCCACACCAGACTCCTCTACT
ATGAGAGCTTCCCTGACCGTGAACCTAAATCTGCTTGGTGAACGGCGATGGTACTGTGAACCTGAAGAGTGCC
TGCAGTGCCAGGGCTGGCAGAGCCGCAGGAGCACCAAGTGTGCTGCAGGAGCTGCCAGGCAGCAGCACATCG
AGATGCTGGCCAACGCCACCCCTGGCTATCTGAAACGTGTGCTCCTGGCCTGACTCTGTGCCACAGGA
CTCTGTGGCTGGCCGTGGACCTGCTGTTGGCTCTGGGCTGTCACTGGCCACCGCTTTGCAAAGTTGTGA
CTCACCAATTCAAGGCCCGAGCTTGACTGTGAAGCATCTGCCATGGGAAGTGTGTTATCCTTCTCT
GTGGCAGTGAAGAAGGAAGAAATGAGAGTCTAGACTCAAGGGACACTGGATGGCAAGAATGCTGCTGATGGTGG
ACTGCTGTGACCTTAGGACTGGCTCCACAGGGTGGACTGGCTGGCCCTGGCTCCAGTCCCTGCCTGGGCCATG
TGTCCCCCTATTCTGTGGCTTTCACTTGCTACTGGCTCTGGCCCTGGGCCCCCGAGCCCTTCCCTATGAGGGATGTT
ACTGGGCTGTGGCTGTGCTACCCAGAGGCTGGGATCGGCTCTGGCCCTGGGTGACCCCTTCCACACACCA
GCCACAGATAGGCTGCCACTGGTCACTGGTAGAGCTGCTGGCTTCCCTGAGCTGGTTCTCGTGGCC
TGACTGGCTTCTGGCGAGCCTAGTAGCTCTGCAGGGCAGTTGTTGCGTTCTCGTGGTTCCAGGC
CCTGGACATCTCACTCCACTCTACCTCCCTAACCAACCAGGAGCATTCAAGCTGGATTGGCAGCAGATGTG
CCCCCAGTCCCGCAGGCTGTTCCAGGGCCCTGATTCTGGATGTGCTATTGGCCCCAGGACTGAAGCTGC
CTCCCCCTCACCTGGGACTGTGGTCCAAGGATGAGAGCAGGGTTGGAGCCATGCCCTCTGGGAACCTATGGA
GAAAGGGAAATCCAAGGAAGCAGCCAAGGCTGCTCGCAGCTCCCTGAGCTGCACCTCTGCTAACCCACCATCA
CACTGCCACCTGCCCTAGGGTCTCACTAGTACCAAGTGGGTCAAGCAGGGCTGAGGATGGGCTCTATCCAC
CCTGGCCAGCACCAAGCTTAGTGTGGACTAGCCCAGAAACTTGAATGGGACCCCTGAGAGAGGCCAGGGTCCCC
TGAGGGCCCCCTAGGGCTTCTGCTGCCAGGGTGTCCATGGATCTCCCTGTTGGCAGCAGGATGGAGAGT
CAGGGCTGCCCTCATGGCAGTAGGCTAAGTGGGTGACTGGCCACAGGCCAGAAAAGGGTACAGCCTCTAGGT
GGGGTCTCAAAGACGCCTCAGGCTGGACTGAGCTGCTCTCCACAGGGTTCTGTCAGCTGGATTCTCTG
TTGCATACATGCTGGCATCTGCTCCCTGTCCTGAGTGGCCCCACATGGGCTCTGAGCAGGCTGTATCTG
GATTCTGGCAATAAAAGTACTCTGGATGCTGTAaaaaaaaaaaaaaaaaaaaaaa

FIGURE 59

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44189
><subunit 1 of 1, 412 aa, 1 stop
><MW: 46658, pI: 6.65, NX(S/T): 4
MGLHLRPYRVGLLPDGLLFLLLLLMLLADPALPAGRHPPVVLVPGDLGNQLEAKLDKPTVVH
YLCSSKKTESYFTIWLNLLELLLPVIIDCWIDNIRLVYNKTSRATQFPDGVDVRVPGFGKTFSL
EFLDPSKSSVGSYFHTMVESLGVWGYTRGEDVRGAPYDWRRAPNENGPYFLALREMIEEMYQ
LYGGPVVLVAHSMGNMYTLYFLQRQPQAWKDKYIRAFVSLGAPWGGVAKTLRVLASGDNNRI
PVIGPLKIREQQRSAVSTSLLPYNTWSPEKVFVQTPTINYTLRDYRKFFQDIGFEDGWLM
RQDTEGLVEATMPPGVQLHCLYGTGVPTPDSFYYESFPDRDPKICFGDGDGTVNLKSALQCQ
AWQSRQEHQVLLQELPGSEHIEMLANATTLAYLKRVLLGP
```

Important features:

Signal peptide:

amino acids 1-28

Potential lipid substrate binding site:

amino acids 147-164

N-glycosylation sites.

amino acids 99-102, 273-276, 289-292 and 398-401

Lipases, serine proteins

amino acids 189-201

Beta-transducin family Trp-Asp repeat

amino acids 353-365

FIGURE 60

CGGACGCGTGGCGGACGCGTGGGCGGCAGCGGCGACGGCGACATGGAGAGCGGG
GCCTACGGCGGGCAAGGCGGGCGCTCCTCGACCTGCGCGCTTCTGACGCAGCGCA
GGTGGTGGCGCGCGCGTGTGCTGGTCTCGCCTGATCGTGTCTCCTGCATCTATGGTG
AGGGCTACAGCAATGCCACGAGTCTAACGAGATGTACTGCGTGTCAACCGAACGAGGAT
GCCTGCCGCTATGGCAGTGCCATCGGGTGCTGGCCTCCTGGCCTCGGCCTCTTCTGGT
GGTCGACGCGTATTCCCCCAGATCAGCAACGCCACTGACCGCAAGTACCTGGTATTGGTG
ACCTGCTCTCTCAGCTCTGGACCTCCTGTGGTTGGTTCTGCTCCTCACCAAC
CAGTGGGCAGTCACCAACCGAAGGACGTGCTGGTGGGGCCGACTCTGTGAGGGCAGCCAT
CACCTTCAGCTTCTTTCCATCTCTGGGTGTGCTGGCCTCCCTGGCCTACCGCGCT
ACAAGGCTGGCGTGGGACGACTTCATCCAGAACATTACGTTGACCCCACCCGGACCCAAACACT
GCCTACGCCCTACCCAGGTGCATCTGTGGACAACCTACCAACAGCCACCCCTCACCCAGAA
CGCGGAGACCACCGAGGGCTACCAGCCCCCCCTGTGTACTGAGTGGCGGTAGCGTGGAA
GGGGGACAGAGAGAGGGCCCTCCCTCTGCCCTGGACTTTCCATCAGCCTCTGGAACGTCCA
GCCCTCTCTTCACCTGTTCCATCCTGTGACACACAGCTAAGGAGCCTCATAGCC
TGGGGGGGCTGGCAGGCCACACCCAAAGTGCCTGTGCCAGAGGGCTTCAGTCAGCCGCT
CACTCCTCCAGGGCACTTTAGGAAAGGGTTTAGCTAGTGTCTCGCTTTAATGA
CCTCAGCCCCGCTGCAGTGGCTAGAACGCCAGCAGGTGCCATGTGCTACTGACAAGTGCCT
CAGCTCCCCCGGCCGGTCAGGCCGTGGAGCCGCTATTATCTGCGTTCTGCCAAAG
ACTCGTGGGGCCATCACACCTGCCCTGTGCAGCGGAGCCGGACCAGGCTTGTGCCCTCA
CTCAGGTTGCTTCCCTGTGCCACTGCTGTATGATCTGGGGCCACCAACCCCTGTGCCGGT
GGCCTCTGGCTGCCCTCCGTGGTGTGAGGGCGGGCTGGTGTCTGGCAGTGGCTTCTG
CTCCCACCCCTGGCAGCAGGAAAGGGCTTGCCTGACAACACCCAGCTTATGTAAATATTC
TGCAGTTGTTACTTAGGAAGCCTGGGAGGGCAGGGTGCCCATGGCTCCAGACTCTGTC
TGTGCCGAGTGTATTATAAAATCGTGGGGAGATGCCGGCTGGATGCTGTTGGAGACG
GAATAATGTTCTCATTCAAAG

FIGURE 61

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48304
<subunit 1 of 1, 224 aa, 1 stop
<MW: 24810, pI: 4.75, NX(S/T): 1
MESGAYGAAKAGGSFDLRRFLTQPQVVARAVCLVFALIVFSCIYGEGYSNAHESKQMYCVFN
RNEDACRYGSAIGVLAFLASAFFLVVDAYFPQISNATDRKYLVIGDLLFSALWTFLWFVGFC
FLTNQWAVTNPKDVLVGADSVRAAITFSFFSIFSWGVLASLAYQRYKAGVDDFIQNYVDPTP
DPNTAYASYPGASVDNYQQPPFTQNAETTEGYQPPPVY
```

Important features:

Type II Transmembrane domain:

amino acids 1-45

Other transmembrane domains:

amino acids 74-90, 108-126 and 145-161

N-glycosylation site.

amino acids 97-100

FIGURE 62

GAGCCACCTACCCCTGCTCCGAGGCCAGGCCCTGCAGGGCCTCATCGGCCAGAGGGTGAATCAGTGACCAGAAGG**ATG**
CCCGTGGCCGAGGCCCTCAGGTGGCTGGCGGGCAGGGGACGGAGGTGATGGCGAGGAAGCGGAGCCAGAGGGG
ATGTTCAAGGCCTGTGAGGACTCCAAGAGAAAAGCCGGGGCTACCTCCGCTGGTCCCCCTGTTGTGCTGCTG
GCCCTGCTCGTGTGGCTCGCGGGGGTGTACTCTGGTATTCCTAGGTACAAGGCGGAGGTGATGGTCAGC
CAGGTGTACTCAGGCAGTCTCGTGTACTCAATGCCACTCTCCCAGGATCTAACCGCCGGGAATCTAGTGC
TTCCGCAGTGAACCGCCAAAGCCAGAAGATGCTCAAGGAGCTCATACCAGCACCGCCCTGGGAACCTACTAC
AACTCCAGCTCGTCTATTCTTTGGGGAGGGACCCCTCACCTGCTTCTCTGGTTCATTCCTCAAATCCCGAG
CACCGCCGGCTGATGCTGAGCCCGAGGTGGTCAGGCAGTGCTGGGGAGGAGCTGCTGCCACAGTCAACAGC
TCGGCTGCCGTCCTACAGGGCGAGTACGAAGTGGACCCCGAGGGCTAGTGATCTGGAAAGCCAGTGTGAAA
GACATAGCTGCATTGAATTCCACGCTGGGTGTTACCGCTACAGCTACGTGGGCCAGGGCCAGGTCTCCGGCTG
AAGGGGCCTGACCACCTGGCCTCCAGCTGCCCTGGCACCTGCAGGGCCCAAGGACCTCATGCTCAAACCTCCGG
CTGGAGTGGACGCTGGCAGAGTGCCGGACCGACTGGCCATGTATGACGTGGCCGGGGCTGGAGAAGAGGCTC
ATCACCTCGGTGTACGGCTGCAGCCGCCAGGAGGCCGTGGGGAGGTCTGGCGTCGGGGCCATCATGGCGGTC
GTCTGGAGAAGGGCCTGCACAGCTACTACGACCCCTTCGTGCTCTCGTGCAGCCGGTGGTCTTCCAGGCCTGT
GAAGTGAACCTGACGCTGGACAACAGGCTGACTCCCAGGGCTCCTCAGCACCCCGTACTTCCCCAGCTACTAC
TCGCCCAAACCCACTGCTCTGGCACCTCACGGTGCCTCTCTGGACTACGGCTTGGCCCTCTGGTTTGATGCC
TATGCACTGAGGAGGCAGAAGTATGATTGGCGTGCACCCAGGGCCAGTGGACGATCCAGAACAGGAGGCTGTG
GGCTTGCGCATCCTGCAGCCCTACGCCAGAGGGATCCCCGTGGTGGCACGGGCCGGGATCACCATCAACTCACC
TCCCAGATCTCCCTCACCGGGCCCGGTGTGCGGGTGCACTATGGCTTGATACAACAGCTGGGACCCCTGCCCTGG
GAGTCCCTCTGTTCTGTGAATGGACTCTGTGTCCTCTGCTGTGATGGGGTCAAGGACTGCCAACGGCCTGGAT
GAGAGAAACTGCGTTGCAAGGCCACATTCCAGTGCAAAGAGGACAGCACATGCACTCCTACTGCCAACGGTCTGT
GATGGGAGGCTGATTGTCCTAACGGCAGCGATGAAGAGGCACTGCCAGGAAGGGGTGCCATGTGGGACATTCA
TTCCAGTGTGAGGACCGAGCTGCGTGAAGAAGCCAACCCCGCAGTGTGATGGGGCCCGACTGCAGGGACGCC
TCGGATGAGGAGCACTGTGACTGTGGCTCCAGGGCCCTCCAGGTTGGGGTGCACACATCTGTGGGGGCCCTCATGCTGACCGC
GGTAGTGGCCATGGCAGGCCAGCCTCCAGGTTGGGGTGCACACATCTGTGGGGGCCCTCATGCTGACCGC
TGGGTGATAACAGCTGCCACTGCTTCCAGGAGGACAGCATGGCCTCACGGTGTGTTGGACCGTGTCCGG
AAGGTGTGGCAGAACTCGCGTGGCTGGAGGGTGTCTTCAGGTGAGCCGCTGCTGCACCCGTACCC
GAAGAGGACAGCCATGACTACGACGTGGCGCTGCTGCAGCTCGACCAACCCGGTGGCGCTCGCCGCGCGTGC
CCCGTCTGCCCTGCCCGCCGCTCCACTTCTCGAGGCCGCTGCACTGCTGGATTACGGCTGGGGGCCCTTG
CGCGAGGGCGGGCCCATCAGCAACGCTCTGCAGAAAGTGGATGTGCAGTTGATCCCACAGGACCTGTGCAGCGAG
GCCTATCGTACCCAGGTGACGCCACGCATGCTGTGTCGGCTACCGCAAGGGCAAGAAGGATGCCCTGTCAGGGT
GAACAGGTGGTCCGCTGGTGTGCAAGGCACTCAGTGGCCGCTGGTTCTGGCGGGCTGGTCAGCTGGGCCCTG
GGCTGTGGCCGGCTAACTACTTCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
ACCTGAGGAACTGCCCTGCAGGCCACCTCTGGACTCAGAGAGGCCAGGGCAACTGCCAAGCAGG
GGGACAAGTATTCTGGCGGGGGTGGGGAGAGAGCAGGCCCTGTGGTGGCAGGGAGGTGGCATCTTGTCTCGTCC
CTGATGTCTGCTCCAGTGTGAGGGCAGGAGGATGGAGAAGTGCCAGCAGCTGGGGGTCAGACGCTCCCTGAGGACC
CAGGCCACACCCAGCCCTCTGCCCTCCAAATTCTCTCTCCGCTCCCTCCACTGCTGCCATAATGCAAG
GCAGTGGCTCAGCAAGAATGCTGGTTCTACATCCCGAGGAGGTGTGAGGTGCCCTGGACTCTGACAGAGG
CTGTTGGGAGGCCCTGCCAGAGAGGCCAGTCCAGCTCGGAAGCCCTGGCTAACTTGGGATCTGGGAAT
GGAAGGTGCTCCCATCGGAGGGGACCCCTCAGAGGCCCTGGAGACTGCCAGGTGGGGCTGCTGCCACTGTAAGCCAA
AAGGTGGGGAGTCCCTGACTCCAGGGTCTTGGCCCACCCCTGCCACCTGGGCCCTCACAGCCCAGACCC
CACTGGGAGGTGAGCTCAGCTGCCCTTGGATAAAGCTGCCCTGATCAAAAAAAAAAAAAAA

40047002 - 402104

FIGURE 63

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49152
><subunit 1 of 1, 802 aa, 1 stop
><MW: 88846, pI: 6.41, NX(S/T): 7
MPVAEAPQVAGGQGDGGDGEAEPEGMFKACEDSKRKARGYLRLVPLFVLLALLVLASAGVL
LWYFLGYKAEVMVSQVYSGSLRVLNRHFSQDLTRRESSAFRSETAKAQKMLKELITTRLGT
YYNSSSVYSFGEGPLTCFFWFILQIPEHRRMLSPVVQALLVEELLSTVNSSAAVPYRAEY
EVDPEGLVILEASVKDIAALNSTLGCYRYSYVGQGQVLRLKPDHLASSCLWHLQGPKDMLM
KLRLWEWTIAECDRDLAMYDVAGPLEKRLITSVYGC SRQEPVVEVLASGAIMAVVWKKGHLHSY
YDPFVLSVQPVVFQACEVNLTLDNRLLSQGVLSTPYFPSYYSPQTHCSWHLTVPSLDYGLAL
WFDAYALRRQKYDLPCTQGQWTIQNRRICGLRILQPYAERIPVVA TAGITINFTSQISLTGP
GVRVHYGLYNQSDPCPGEFLCSVNGLCPACDGVKDCPNGLDERNCVCRATFQCKEDSTCIS
LPKVCDGQPDCLNGSDEEQCQEGVPCGTFQCEDRSCVKKPNCQCDGRPDCRDGSDEEHCD
CGLQGPSSRIVGGAVSSEGEWPWQASLQVRGRHICGGALIADR WVITAHCQEDSMASTVL
WTVFLGKVWQNSRWPGEVSFKVSRLLLHPYHEEDSHDYDVALQLDHPVVRSAAVRPVCLPA
RSHFFEPGLHCWITGWGALREGGPISNALQKVDVQLIPQDLCSEAYRYQVTPRMLCAGYRKG
KKDACQGDGGPLVCKALSGRWFLAGLVSWGLGCGRPNYFGVYTRITGVISWIQQVVT
```

Important features:

Type II transmembrane domain:

amino acids 46-67

Serine proteases, trypsin family, histidine active site.

amino acids 604-609

N-glycosylation sites.

amino acids 127-130, 175-178, 207-210, 329-332, 424-427, 444-447
and 509-512

Kringle domains.

amino acids 746-758 and 592-609

Homologous region to Kallikrein Light Chain:

amino acids 568-779

Homologous region to Low-density lipoprotein receptor:

amino acids 451-567

FIGURE 64

GCACCCAGGCCAGTGGACGATCCAGAACAGGAGGCTGTGGCTTGCATCCTGCAGCCC
TACGCCGAGAGGATCCCCGTGGCGACGCCGGGATCACCACAACTCACCTCCCAGAT
CTCCCTCACCGGCCGGTGTGCGGGTGCACATGGCTTGTACAACCAGTCGGACCCCTGCC
CTGGAGAGTTCTCTGTGAATGGACTCTGTGTCCCTGCCTGTGATGGGTCAAGGAC
TGCCCCAACGGCTGGATGAGAGAACTGCAGTTGCAGAGCCACATTCCAGTGCAAAGAGGA
CAGCACATGCATCTCACTGCCAAGGTCTGTGATGGCAGCCTGATTGTCTAACGGCAGCG
ATGAAGAGCAGTGCAGGAAGGGTGCATGTGGACATTCACCTCAGTGTGAGGACCGG
AGCTGCGTGAAGAACGCCAACCGCAGTGTGATGGCGGCCGACTGCAGGGACGGCTCGGA
TGAGGAGCACTGTGACTGTGGCCTCCAGGGCCCTCCAGCCGCATTGTTGGTGGAGCTGTGT
CCTCCGAGGGTGAGTGGCCATGGCAGGCCAGCCTCCAGGTTGGGTCGACACATCTGTGG
GGGCCCTCATCGCTGACCGCTGGTGATAACAGCTGCCACTGCTTCCAGGAGGACAGCAT
GGCCTCCACGGTGTGGACCGTGTGGCCTGGCAAGGTGTGGCAGAACTCGCGCTGGCCTG
GAGAGGTGTCTTCAAGGTGAGGCCCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGTGCGCCC
GACTACGACGTGGCGCTGCTGCAGCTCGACCACCCGGTGGTGCCTCGGCCGTGCGCCC
CGTCTGCCTGCCCGCGCTCCACTTCTCGAGGCCGGCCTGCACTGCTGGATTACGGCT
GGGCGCCTTGCAGGGCGGCCATCAGCAACGCTTGCAGAAAGTGGATGTGCAGTTG
ATCCCACAGGACCTGTGCAAGGCCCTATCGCTACCGAGGTGACGCCACGCATGCTGTGTGC
CGGCTACCGCAAGGGCAAGAAGGATGCCTGTCAAGGTGACTCAGGTGGTCCGCTGGTGTGCA
AGGCACACTGGCCGCTGGTTCTGGCGGGCTGGTCAGCTGGGCTGGCTGTGGCCGG
CCTAACTACTTCCGGCGTCTACACCCGCATCACAGGTGTGATCAGCTGGATCCAGCAAGTGGT
GACCTGAGGAACGTGCCCTGCAAAGCAGGGCCCACCTCCTGGACTCAGAGAGCCAGGGC
AACTGCCAAGCAGGGGACAAGTAT

FIGURE 65

GGACGAGGGCAGATCTCGTTCTGGGGCAAGCCGTTGACACTCGCTCCCTGCCACCGCCCCGGG
CTCCGTGCCGCCAAGTTTCATTTCACCTTCTGCCCTCCAGTCCCCCAGCCCCCTGGCCG
AGAGAAGGGTCTTACCGGCCGGATTGCTGGAAACACCAAGAGGTGGTTTTGTTTTAAA
ACTTCTGTTCTGGGAGGGGGTGTGGCGGGGCAGGATGAGCAACTCCGTTCCCTGCTCTG
TTTCTGGAGCCTCTGCTATTGCTTTGCTGCGGGGAGGCCCGTACCTTTGGTCCAGAGGGAC
GGCTGGAAGATAAGCTCCACAAACCCAAAGCTACACAGACTGAGGTCAAACCATCTGTGAGG
TTAACCTCCGACCTCCAAGGACCCAGAGCATGAAGGATGCTACCTCTCCGTCGGCACAG
CCAGCCCTAGAAGACTGCAGTTCAACATGACAGCTAAACCTTTCATCATTACGGAT
GGACGATGAGCGGTATCTTGAAAAGCTGGCTGCACAAACTCGTGTCAAGCCCTGCACACAAGA
GAGAAAGACGCCAATGTAGTTGTGGTTGACTGGCTCCCCCTGGCCACCAAGCTTACACGGA
TGCGGTCAATAATACCAGGGTGGTGGACACAGCATTGCCAGGATGCTGACTGGTCAGG
AGAAGGACGATTTCTCTCGGAATGTCCACTTGATGGCTACAGCCTCGAGCGCACGTG
GCCGGTATGCAGGCAACTCGTGAAGGAACGGTGGCCGAATCACAGGTTGGATCCTGC
CGGGCCATGTTGAAGGGGCCGACATCCACAAGAGGCTCTCCGGACGATGCAGATTTG
TGGATGTCCTCCACACCTACACGCGTTCTCGCTTGAGCATTGGTATTCAAGATGCTGTG
GGCCACATTGACATCTACCCCAATGGGGGTGACTCCAGCCAGGCTGTGGACTCAACGATGT
CTTGGGATCAATTGCAATGCAATGGAACAATCACAGAGGTGGTAAAATGTGAGCATGAGCGAGCCG
TCCACCTCTTGACTCTCTGGTGAATCAGGACAAGCCGAGTTTGCTCCAGTCAGTC
GACTCCAATCGCTTCAAAAGGGATCTGTCTGAGCTGCCGAAGAACCGTTGAATAGCAT
TGGCTACAATGCCAAGAAAATGAGGAACAAGAGGAACAGCAAAATGTACCTAAAAACCCGGG
CAGGCATGCCTTCAGAGGTAACCTCAGTCCCTGGAGTGTCCCTGAGGAAGGCCCTTAATA
CCTCCTCTTAATACCATGCTGCAGAGCAGGGCACATCCTAGCCCAGGAGAAGTGGCAGCA
CAATCCAATCAAATCGTTGCAAATCAGATTACACTGTGCATGTCTAGGAAAGGAATCTT
ACAAAATAAACAGTGTGGACCCCTAATAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAA

FIGURE 66

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49646
><subunit 1 of 1, 354 aa, 1 stop
><MW: 39362, pi: 8.35, NX(S/T): 2
MSNSVPLLCFWSLCYCFAAGSPVPFGPEGRLEDKLHKPKATQTEVKPSVRFLRTSKDPEHE
GCYLSVGHSQLPLEDCSFNMTAKTFFIIHGWTMSGIFENWLHKLV SALHTREKDANVVVDWL
PLAHQLYTDAVNNTRVVGHSIARMLDWLQEKKDDFSLGNVHLIGYSLGAHVAGYAGNFVKGT
GRITGLDPAGPMFEGADIHKRLSPDDADFVDVLHTYTRSFGLSIGIQMPVGHIDIPNGGDF
QPGCGLNDVLGSIAYGTITEVVKCEHERAVHLFVDSLNVQDKPSFAFQCTDSNRFKKGICL
CRKNRCNSIGYNAKKMRNKRNSKMYLKTRAGMPFRGNLQSLECP
```

Important features:

Signal peptide:

amino acids 1-16

Lipases, serine active site.

amino acids 163-172

N-glycosylation sites.

amino acids 80-83 and 136-139

FIGURE 67

CGGACGCGTGGCGGACGCCCTGGGCAAGGGCGGGCGGGCGAGCCACCTCTTCCCTCCCCCGC
TTCCCTGCGCTCGCTGGACCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCTGGGGCTGACAGT
CGGAAAGTTGGCCCAGAGGAAGTGGTCTAAACCCCGCAGGTGGCGACCAGGGCGCTCG
CTGCGCTGCGGGCGGGCTGTAGGCAGGGCGCCCGAGTGCAGAGACCCGGGCTCAGGAGCGGCCCCGGAG
AGAAGAGTGCAGGGCGGAGAAAACAACCTCAAAGTTGGCGAAAGGCACCCGGCTACTCCCGGGCTGCCG
CCGCGTCCCCGCCCCAGCCCTGGCATCCAGAGTACGGGTGAGGCCCATGGAGCCCCCTGGGGAGGCGG
CACCAAGGGAGCCTGGCGCCGGCTCCGCGCAGCCCCTGGGTAGACCACAGAAGCTCCGGACCTTCCG
GCACCTCTGGACAGCCAGGATGCTGGCACCCCTCTCTCTCTGGAGGCGCTGGCCATCCAG
ACCGGATTATTTCAAATCATGCTTGTGAGGACCCCCAGCAGTGTCTTAAAGTGCAGGGCACCTAACAGA
GGCCCTGGTCGGGACAGCCGACCTCCCTGCGAACCTGCGACCTGGCTCATCTGGCAGAAGGAACAGACTG
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCCCTACGCTCCCTCTCAGCCAC
TGATCTCCCTGTGTGAGGCACCTCCAGCCCTCTGCAGCTGCCGGGGCAACGTCACCATCACTAACAGCTATG
CTGGGCCAGAGCACCCATGGCCAGGGCTCTGCTCTACAGCCAAGATTGGCTGATGTGCCGTGAGGAAG
AGTTCACTGCGCTGAACCACCGCTGTGTATCTGCTGCCAGCGCTGTGATGGGTTGATGCTGTGGCGATGGCT
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCTGGCCTGACCCCAAGACCCGTCCTCCCTGCCATGCAATG
TCACCTTGAGGACTTCTATGGGTCTCTCTGGATATACACACCTAGCCTCAGTCTCCACCCACTTCA
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCTGGCGCTTCACAGCCCTGGACTTGGCTTGG
GAGATGCACTGCACTGTGTATGACGGCCCTGGGCCCCCTGAGAGCTCCGACTACTGCGTAGTCTACCCACTTCA
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTGTGCTTACACACAGTTGCTGGAGCA
ATGGTCGTGGCTTAATGCCACCTACATGTGGGGCTATTGCTGCCCTGGACAGACCTGTGGCTTAGGCT
CTGGCTGGAGGCTGGCAAGGCTAGGTGAGCGCTGCTACAGTGAGGACAGCGCTGTGACGGCTATGGACT
GTGCTGACGGCACAGATGAGGAGACTGCCACGGCTGCCACCTGGACACTTCCCTGTGGCTGGCACCT
CTGGTGCACAGCCTGCTACCTGCTGACCGCTGCAACTACAGACTTTCTGTGCTGATGGAGCAGATGAGA
GACGCTGTGGCATTGCCAGCTGGCAATTCCGATGGGGAGGAGAAGTGCCTGATGAGACGTGGGTGTGCG
ATGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGACTGCTCTATGTTCTGCCCAAGGTCTTACAG
CTGCACTGCACTGGCAGCTAGTGTGCGGCTGCTCTGGTCACTGCCCTGGCTGACACTGCAAGCTCTATGCCA
TTCGCACCCAGGAGTACAGCATCTTGCCCCCTCTCCGGATGGAGGCTGAGATTGTGCAAGCAGGACCCCC
CTTCCTACGGGAGCTCATTGCCAGGGTGCATCCCACCTGTAGAACACTTCTACAGAGAACCTTAATGATA
ACTCAGTGCTGGCAACCTGCGTTCTGTCTACAGATCTACGCCAGGATATGACTCCAGGAGGTGGCCAGGTG
CCCGCGTCGTCAAGGGGGCGCTTGATGCGACGCCCTGGTACGCCCTGCCCTGGCTCCCTCGAA
CCAACACCCGGCTCGGGCTCTGAGGCCAGATCCCAGGTACACCTCTGCTGCTCCCTTGAGGCCCTAGATG
GTGGCACAGGTTCCAGCCGTAGGGGGCAGTGGGTGGCAAGATGGGAGCAGGACCCCCACTGCCCATCA
AGGCTCCCTCCATCTGCTAGCACGTCTCCAGCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC
TGCCCTAGAGGCATCACTATTGTCTGGAGTGGTGCAGGCCCTGGAGGCCCTGGCTCCCTGGGGCCCC
CAGGACCAACCGGAGGCCCTGGACCCACACAGCAGTCTGGCCCTGGAGATGAGGACGATGTGCTACTGG
TGCCACTGGCTGAGCCGGGGGTGTGGTAGCTGAGGCAGAGGATGAGGCCACTGCTTACCTGAGGGACCTGGGG
CTCTACTGAGGCTCTCCCTGGGGCTACTCATAGTGTGCAACACTTTAGAGGTGGTCAAGCTCCCTCC
ACCACTCCCTCTGCTCCCTGGATTTCAGGGACTTGGTGGCCCTCCGTTGACCTATGTAGCTGCTATAAAGT
TAAGTGTCCCTCAGGCAGGGAGAGGGCTACAGAGTCTCTGCTACGTGGCCATGGCCAGACACCCCCAGTCCCT
TCACCACTGCTCCCAAGCCACCCATTGGGTGGCTGTTTTAAAAAGTAAAGTCTTAGAGGATCATA
GGCTGGACACTCCATCTTGCCTAACCTCTACCCAAAGTGGCTTAAGCACCGGAATGCCAATTAACAGAGA
CCCTCCAGCCCCAAGGGAGGATTGGGAGAACCTGAGGTTTGCCATCCACAATCCCTCTACAGGGCCTGG
CTCACAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAATAAA
GGAATCATACTCTC

FIGURE 68

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631
<subunit 1 of 1, 713 aa, 1 stop
<MW: 76193, pI: 5.42, NX(S/T): 4
MLLATLLLLLGGALAHPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLIL
GSKEQTVTIRFQKLHLACGSERLTLRSPLOPLISLCEAPPSPLQLPGGNVTITYSYAGARAP
MGQGFLLSYSQDWLMCLQEEFQCLNHRCVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRP
VPSLPCNVTLLEDFYGVFSSPGYTHLASVSHPQSCHWLLDPHDGRRLAVRTALDLGFGDAVH
VYDGPGPPESSRLLRSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRFNATYHVRGYCLP
WDRPCGLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPPGHFPCGAAGTSGAT
ACYLPADRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPDCAFGSDEWDCS
YVLPRKVITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSY
GQLIAQGAIPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGPGARRRQRGRILMRRILVR
RLRRWGLLPRNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKA
PLPSASTSPAPTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLLPSLGPPGPTRSPPGPHTAV
LALEDEDDVLLVPLAEPGVWVAEAEDEPLLT
```

Important features:

Signal peptide:

amino acids 1-16

Transmembrane domain:

amino acids 442-462

LDL-receptor class A (LDLRA) domain proteins

amino acids 411-431, 152-171, 331-350 and 374-393

FIGURE 69

CGAGCTGGCGAGAAGTAGGGAGGGCGGTGCTCCGCCCGGTGGCGGTTGCTATCGCTTCG
CAGAACCTACTCAGGCAGCCAGCTGAGAAGAGTTGAGGGAAAGTGCCTGCTGGGCTGCA
GACGCGATGGAATAACGTGCAGCCGAAAATAAAACATCGCCCCCTCTGCTTCAGTGTGAAAGG
CCACGTGAAGATGCTGCGGCTGGCACTAACTGTGACATCTATGACCTTTTATCATCGCAC
AAGCCCCTGAACCATAATTGTTATCACTGGATTGAAGTCACCGTTATCTTATTTTCATA
CTTTTATATGTAUACTCAGACTTGATCGATTAATGAAGTGGTTATTTGGCCTTGCTTGATAT
TATCAACTCACTGGTAACAAACAGTATTGTCATCGTATCTGTGTTGGCACTGATACCAAG
AAACCCACAAACATTGACAGTTGGTGGAGGGGTGTTGCACCTGTGACAGCAGTATGCTGTCTT
GCCGACGGGCCCTTATTTACCGGAAGCTCTGTTCAATCCCAGCGGTCTTACCAAGAAAAAA
GCCTGTGCATGAAAAAAAAGAAGTTTGTAATTTATATTACTTTAGTTGATACTAAGT
ATTAACATATTCTGTATTCTCCAAAAA

FIGURE 70

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49645
><subunit 1 of 1, 152 aa, 1 stop
><MW: 17170, pI: 9.62, NX(S/T): 1
MDNVQPKIKHRPFCFSVKGHVKMLRLALTTSMTFFIIAQAPEPYIVITGFEVTVILFFILL
YVLRLDRLMKWLFWPLLDIINSLVTTVFMLIVSVLALIPETTTLTVGGGVFALVTAVCCLAD
GALIYRKLLFNPSGPYQKKPVHEKKEVL
```

Important features:

Potential type II transmembrane domain:
amino acids 26-42

Other potential transmembrane domain:
amino acids 44-65, 81-101 and 109-129

Leucine zipper pattern
amino acids 78-99 and 85-106

N-myristoylation site.
amino acids 110-115

Ribonucleotide reductase large subunit protein
amino acids 116-127

FIGURE 71

GGCGAGAAGTAGGGAGGGCGTGTCCGCCGCGTGGCGGTTGCTATCGTTTGACAGAAC
TACTCAGGCAGCCAGNTGAGAAGAGTTGAGGGAAAGTGCTGCTGCTGGCTGCAGACGCGA
TGGATAACGTGCAGCCGAAAATAAAACATCGCCCTTCTGCTTCAGTGTGAAAGGCCACGTG
AAGATGCTGCGGCTGGCACTAAGTGNGACATCTATGACCTTTTATNATCGCACAAAGCCCC
TGAACCATAATTGTTATCACTGGATTGAAAGTCACCGTTATCTTATTTCATACTTTAT
ATGTAACACTGATCGATTAATGAAGTGGTATTTGGCCTTGCTTGATATTATCACAC
TCACTGGTAACAAACAGTATTGCTCATCGTATCTGTGTTGGCACTGATACCAGAAACCAC
AACATTGACAGTTGGTGGAGGGTGTGCACTGTGACAGCAGTATGCTGTNTGCCGAC

FIGURE 72

CAGCCCCGGCGCCGGCCGAGTCGCTGAGCCGGCTGCCGGACGGGACGGGACCGGCTAGG
CTGGGCGCCTCCGGGGCCGTGGCATGGCGCACTGGCCGGCTGCTGCTGC
CTCTGCTGGCCCAGTGGCTCTGCGCCGCCGGAGCTGGCCCCGGCCCTTACGCTG
CCCCCTCCGGGTGGCCGCCAGAACCGCGTAGTTGCGCCCACCCGGACCCGGACCCCG
TGCCGAGGCCACGCCGACGGCTTGGCGCTGCCCTGGAGCCTGCCCTGGCGTCCCCGG
GCGCCGCCAACTTCTGGCCATGGTAGACAACCTGCAGGGGACTCTGGCCGGCTACTAC
CTGGAGATGCTGATCGGGACCCCCCGCAGAACGCTACAGATTCTCGTTGACACTGGAAGCAG
TAACTTTGCCGTGGCAGGAACCCCGACTCCTACATAGACACGTACTTGACACAGAGAGGT
CTAGCACATACCGCTCCAAGGGCTTGACGTACAGTGAAGTACACACAAGGAAGCTGGACG
GGCTTCGTTGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTCTGTCAA
CATTGCCACTATTTGAATCAGAGAATTCTTTGCCCTGGATTAAATGGAATGGAATAC
TTGGCCTAGCTTATGCCACACTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTCGACTCC
CTGGTACACAAGCAAACATCCCCAACGTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCC
CGTTGCTGGATCTGGGACCAACGGAGGTAGTCTGTCTGGGTGGAATTGAACCAAGTTGT
ATAAAGGAGACATCTGGTATAACCCCTATTAGGAAGAGTGGTACTACCAGATAGAAATTCTG
AAATTGGAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGC
CATCGTGGACAGTGGCACCAACGCTGCGCTGCCCTGGAGGTGTTGATGCCGTGG
AAGCTGTGGCCCGCGCATCTGATTCCAGAATTCTCTGATGGTTCTGGACTGGTCCCAG
CTGGCGTGGACGAATTGGAAACACCTGGCTTACTCCCTAAAATCTCCATCTACCT
GAGAGACGAGAACTCCAGCAGGTATTCCGTATCACAACTGCCCTCAGCTTACATTAGC
CCATGATGGGGCCGGCCTGAATTATGAATGTTACCGATTGGCATTCCCCATCCACAAAT
GCGCTGGTATGGTGCACGGTGATGGAGGGCTTACGTACATCTCGACAGAGCCAGAA
GAGGGTGGCTTCGCAGCGAGCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATT
CCGGGCCTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTGAGCGAG
CCCATTGTTGGATTGTGTCTATGCCCTCATGAGCGTCTGTGGAGCCATCCTCTGTCTT
AATCGTCCTGCTGCTGCCGTTCCGGTGTCAAGCGTCGCCCCGTGACCGTGGAGGTCA
ATGATGAGTCCTCTGGTCAGACATCGCTGGAAATGAATAGCCAGGCTGACCTCAAGCAA
CCATGAACTCAGCTATTAGAAAATCACATTCCAGGGCAGCAGCCGGATCGATGGTGGCG
CTTCTCCTGTGCCAACCGCTTCAATCTCTGTTCTGCTCCCAGATGCCTTAGATTAC
TGTCTTTGATTCTGATTTCAAGCTTCAAATCCTCCCTACTTCCAAGAAAAATAATTAA
AAAAAAACTTCATTCTAA

FIGURE 73

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pi: 5.08, NX(S/T): 2
MGALARALLPLLAQWLLRAAPELAPAPFTLPLRVAATNRVVAPTPGPGTPAERHADGLAL
ALEPALASPAGAANFLAMVDNLQGDSGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHS
YIDTYFDTERSSTYRSKGFDVTVKYTQGSWTGFVGEDLVTIPKGFNTSFLVNIATIFESENF
FLPGIKWNGILGLAYATLAKPSSSLETFDSLVTQANIPNVFSMQMCAGLPVAGSGTNGGS
LVLGGIEPSLYKGDIWYTPIKEEWYYQIEILKLEIGGQSLNLDREYNADKAIVDSGTTLLR
LPQKVFDAVVEAVARASLIPEFSDGFWTGSQIACWTNSETPWSYFPKISIYLRDENSSRSFR
ITILPQLYIOPMMGAGLNYYECYRGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCA
EIAGAAVSEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLVLIVLLLLPFRC
QRRPRDPEVVNDESSLVRHRWK

Important features:

Signal peptide:

amino acids 1-20

Transmembrane domain:

amino acids 466-494

N-glycosylation sites.

amino acids 170-173 and 366-369

Leucine zipper pattern.

amino acids 10-31 and 197-118

Eukaryotic and viral aspartyl proteases

amino acids 109-118, 252-261 and 298-310

FIGURE 74

CGCCTCCGCCCTCGGAGGCTGACGCCCGGGCGCCGTTCCAGGCCTGTGCAGGGCGGATCG
GCAGCCGCCTGGCGCGATCCAGGGCGGTGCAGGGCGCTGGCGGGAGCCGGAGGCCGCC
GGCATGGAGGCCTGCTGCTGGCGCGGGTTGCTGCTGGCGCTTACGTGCTTGTCTACTA
CAACCTGGTGAAGGCCCGCCGTGCGCGGCATGGCAACCTGCGGGGCCACGGCGTGG
TCACGGCGCCAACAGCGGCATCGGAAAGATGACGGCGCTGGAGCTGGCGGCCGGGAGCG
CGCGTGGTGCCTGGCGCCAGCCAGGAGCGCGGGAGGCAGCTGCCTCGACCTCCGCA
GGAGAGTGGAAACAATGAGGTATCTCATGGCCTTGGACTTGGCAGTCTGGCTCGGTGC
GGGCCTTGCCTGCCTTGAGCTTGAGCCACGGTTGGACATCCTCATCCACAATGCC
GGTATCAGTTCTGTGGCCGGACCGTGAGGCCTTAACCTGCTGCTTGGTGAACCATAT
CGGTCCCTTCTGCTGACACATCTGCTGCTGCCTGCGCTGAAGGCATGTGCCCTAGCCGCG
TGGTGGTGGTAGCCTCAGCTGCCACTGTCGGGGACGTCTGACTTCAAACGCCCTGGACCGC
CCAGTGGTGGCTGGCGCAGGAGCTGCCAACAGCTTGAAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
GTTTGCCTGGGAGCTGCCAACAGCTTGAAGGCCACTGGCGTCACCTGCTATGCAGCCCACC
CAGGGCCTGTGAACTCGGAGCTGTTCTGCCATGTTCTGGATGGCTGCCACTTTG
CGCCCATTGGCTTGGCTGGTGCTCCGGCACCAAGAGGGGTGCCAGACACCCCTGTATTG
TGCTCTACAAGAGGGCATCGAGCCCTCAGTGGAGATATTTGCCAACTGCCATGTGGAAG
AGGTGCCCTCAGCTGCCAGACGACCGGGCAGCCCATCGGCTATGGAGGCCAGCAAGAGG
CTGGCAGGGCTTGGGCTGGGAGGATGCTGAACCCGATGAAGACCCCCAGTCTGAGGACTC
AGAGGCCCATCTCTCTAACGACCCCCCACCCTGAGGAGCCCACAGTTCTAACCTTACC
CCAGCCCTCAGAGCTCACCAGATTTGCTAACGATGACGACCGAACATTGAGCTAAAGTTGAG
CCTGAGATCCAGCTCTCCTTAACCCTCAGGCCAGGATGCTGCCATGGCACTTCATGGCCTT
GAAAACCTGGATGTGTGAGGCCATGCCCTGGACACTGACGGTTGTGATCTTGACCTC
CGTGGTTACTTCTGGGCCCCAAGCTGTGCCCTGGACATCTCTTCTGGTGAAGGAAT
AATGGGTGATTATTCCTCTGAGAGTGTGACAGTAACCCAGATGGAGAGATAGGGTATGCT
AGACACTGTGCTCTCGGAAATTGGATGTAGTATTTCAGGCCACCCCTATTGATTCTG
ATCAGCTCTGGAGCAGGGCAGGGAGTTGCAATGTGATGCACTGCCAACATTGAGAATTAG
TGAACGTACCCCTTGCAACCGTCTAGCTAGGTAGTTAAATTACCCCATGTTAATGAAGCG
GAATTAGGCTCCCGAGCTAAGGGACTGCCCTAGGGTCTCACAGTGAGTAGGAGGAGGGCCTG
GGATCTGAACCCAAGGGTCTGAGGCCAGGGCGACTGCCGTAAAGATGGGTGCTGAGAAGTGA
GTCAGGGCAGGGCAGCTGGTATCGAGGTGCCCATGGAGTAAGGGACGCCCTCCGGCG
ATGCAGGGCTGGGTCATCTGTATCTGAAGCCCTCGGAATAAGCGCGTTGACCGCAAAA
AAAAAAAAAAAAAAA

40017082-1402404

FIGURE 75

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48227
<subunit 1 of 1, 377 aa, 1 stop
<MW: 40849, pI: 7.98, NX(S/T): 0
MEALLLGAGLLLGAYVLVYYNLVKAPPCCGGMGNLRGRTAVVTGANSIGKMTALELARRGAR
VVLACRSQERGEAAAFDLRQESGNNEVIFMALDLASLASVRAFATAFLSSEPRLDILIHNAG
ISSCGRTREAFNLLLRVNHIGPFLTHLLLPCLKACAPSRRVVVASAHHCRGRLLDKRLDRP
VVGWRQELRAYADTKLANVLFARRELANQLEATGVTCYAAHPGPVNSELFRLHVPGWLRPLL
PLAWLVLRAPRGGAQTPLYCALQEGIEPLSGRYFANCHVEEVPPAARDDRAAHRLWEASKRL
AGLGPGEDAEPDEDPQSEDSEAPSSLSTPHPEEPTVSQPYPSHQSSPDLSKMTHRIQAKVEP
EIQLS
```

Important features:

Signal peptide:

amino acids 1-16

Glycosaminoglycan attachment site.

amino acids 46-49

Short-chain alcohol dehydrogenase family

amino acids 37-49 and 114-124

FIGURE 76

GGAGGAGACAGCCTCCTGGGGGGAGGGGTCCTGCCTCTGCTCCTGCTCATCATGGGAGGCATGGCTCAG
GACTCCCCGCCCCAGATCCTAGTCCACCCCCAGGACCAGCTGTTCCAGGGCCCTGGCCAGGATGAGCTGC
CAAGCCTCAGGCCAGCCACCTCCCACCATCCGCTGGTGTGAATGGCAGCCCTGAGCATGGTGCCAGC
CCACACCACCTCCTGCCTGATGGGACCCCTCTGCTGCTACAGCCCCCTGCCCAGGACATGCCACGATGGCCAG
GCCCTGTCCACAGACCTGGGTGTCACACATGTGAGGCCAGCAACGGGCTTGGCACGGCAGTCAGCAGAGGCGCT
CGGCTGTCTGGCTGTCTCCGGGAGGATTTCAGATCCAGCCTCGGGACATGGTGGCTGTGGTGGTGAGCAG
TTTACTCTGGAATGTGGGCCCTGGGCCACCCAGGCCACAGTCTCATGGTGAAGGAAAGATGGAAACCCCTG
GCCCTCCAGCCCGAAGGCACACAGTGTCCGGGGTCCCTGCTGATGGCAAGAGCAGAGAAGAGTGAAGGG
ACCTACATGTGTGGCCACCAACAGCGCAGGACATAGGGAGAGCCGGCAGCCGGTTTCCATCCAGGAGCCC
CAGGACTACACGGAGCCTGGAGCTTCTGGCTGTGCAATTCACTGGAAAATGTGACACTGCTGAACCCGGAT
CCTGCAGAGGGCCCAAGCCTAGACCGGGCTGTGGCTCAGCTGGAAGGTCACTGGCCCTGCTGCGCTGCCAA
TCTTACACGGCCTGTTCAAGGACCCAGACTGCCCGGGAGGCCAGGGAGCTCCGTGGGAGGACTGCTGGCC
GGCTGGCAGAGCGCAGAGCTGGAGGCCACTGGGCAAGACTACGAGTTCAAAGTGAGGACATCCTCTGGC
CGGGCTCGAGGCCCTGACAGCAACGTGCTGCTCTGAGGCTGCCGGAAAAAGTGCCAGTGCCACCTCAGGAA
GTGACTCTAAAGCCTGGCAATGGCACTGTCTTGTGAGCTGGTCCCACCTGCTGAAAACCACAATGGCATH
ATCCGTGGCTACCAGGTCTGGAGCCTGGCAACACATCACTGCCACCAGCAACTGGACTGTAGTTGGTGAGCAG
ACCCAGCTGAAATGCCACCCATATGCCAGGCTCTACTGCGTGAAGTGGCTGCACTGGTGTGAGGAGCT
GGGGAGGCCAGTAGACCTGCTGCTCTTTAGAGCAGGCCATGGAGCAGGCCACCAAGAACCCAGTGAGCAT
GGTCCCTGGACCCCTGGAGCAGCTGAGGGCTACCTTGAGCGGCCCTGAGGTCAATTGCCACCTGCGGTGTTGCACTC
TGGCTGCTCTCTGGGACCCGGCTGTGTATCCACCGCCGGCGAGCTAGGGTGCACCTGGGCCAGGTCTG
TACAGATATAACCACTGAGGATGCCATCCTAAAACACAGGATGGATCACAGTGACTCCCAGTGGTGTGAGACACT
TGGCGTCCACCTCTGGCTCTCGGGACCTGAGCAGCAGCAGCAGCCTCAGCAGTGGCTGGGGCGGATGCCGG
GACCACTAGACTGTCGTCGCTCTTGCTCTCTGGGACTCCCGAAGCCCCGGCTGCCCTGCTTCCAGACACC
AGCACTTTTATGGCTCCCTCATCGTGAAGCTGCCCTCCAGCTCACAGCAGTGGCTCCAGCTCAGACGCCCTGCA
GAGGCCAGTCTCTCCCCGCTTGTCTCTGGCCCTGAGGCCCTGGGAAAGGCAAGTCCCAGGTCCAGCT
GCCAACAGTCTCCCACTGTCGCCGGCAGCCACTCTGGAGCTCCGGGCTGTGAGTTAGGAAATAGAGGTTCC
AAGAACCTTCCCAAAGGCCAGGAGCTGTGCCCAAGCTCTGGTGCCTGGGGGCTGGGACCGAAACTCCTC
AGCTCCTCAAATGAGCTGGTTACTCGTCATCTCCCTCCAGCACCCCTTTCTCATGAAACTCCCCAACTCAG
AGTCAACAGACCCAGCCTCCGGTGGCACACAGGCTCCCTCCATCTGCTGCCAGCAGCCCCATCCCCATC
CTTAGCCCCCTGAGTCCCCCTAGCCCCCAGGCCCTTCCCTCTGGCCCCAGGCCAGTCCAGTCGCTGTCC
AGCTCCTCACTGTCATCCCTGGGGAGGATCAAGACAGCGTGTGACCCCTGAGGAGGTAGCCCTGTGCTTGGAA
CTCAGTGGGTGAGGAGACTCCAGGAACAGCGTCTCTCCATGCCAAGGGCTCTTACCCCCCACCACCTAT
GGGTACATCAGCGTCCCAAAGCCTCAGAGTTACGGACATGGGAGGACTGGAGGAGGGTGGGGCCCAAGGG
GGAGTCTTGTGTGCCACCTCGGCCCTGCCACCCCCACCCCCAGCGAGGGCTCTTAGCCAATGGTTGGGC
TCAGCCCTCTGAGGAAATGCCGCCAGGCCAGGCCAGGCCAGCTTGTGAGCTCCTCCGATGGCTCTTCCTGCTGAT
GCTCACTTTGCCGGGCCCTGGCAGTGGCTGTGAGTAGCTTGGTTCTGGCTAGAGGCCAGGGAGGCAGACTGC
GTCTTACATAGATGCCCTCATCACCTCCCTCCCCACGGGATGAGATCTCTGACCCCCAACCTCTCCCTGCCCTG
TGGGAGTGGAGGCCAGACTGGTGGAGACATGGAGGTCAGCCACACCCAGCGGCCAGGGATGCCCTCC
TGGCCCCCTGACTCTCAGATCTCTCCCTAGAGAAGCTAGCTCCAGCTCCACTGCGTATGCCCAAGGCTGGTGTCTCT
GTAGATTACTCCTGAACCGTGTCCCTGAGACTTCCCAGAGGGAAATCAGAACCAACTCTCTGTCCACCCACAAG
ACCTGGGCTGTGGTGTGGCTTGTGTTCTCTGAGCTGGGCTCAGCTGGGAGGCTCTTCCACCTCCAGGAG
TTCTCCCTCCACGATTGTGAAAACAATGAAAACAAAATTAGAGCAAAGCTGACCTGGAGCCCTCAGGGAGCAA
ACATCATCTCCACCTGACTCTAGCCACTGCTTCTCTGCCCCATCCACTCCACCAACCCAGGTTGTTGGC
CTGAGGAGCAGCCCTGCCCTGCTGCTCTCCCCACCATTTGGATCACAGGAAGTGGAGGAGGCCAGGGTGCCTT
GTGGAGGACAGCAGTGGCTGTGGAGAGGGCTGTGGAGGAAGGAGCTCTCGGAGGCCCTCTCAGCCTTACCT
GGGCCCTCTAGAGAAGAGCTCAACTCTCTCCACCATGGAAAGAAAATAATTATGAATGCCACTG
AGGCACGTGAGGCCCTACCTCATGCCAAACAAAGGGTCAAGGCTGGTCTAGCGAGGATGCTGAAGGAAGGGAGG
TATGAGACCGTAGGTCAAAGCACCACCTCGTACTGTGCACTATGAGCTTAAGAAATTGATACCATAAAAT
GGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 77

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41404
<subunit 1 of 1, 985 aa, 1 stop
<MW: 105336, pI: 6.55, NX(S/T): 7
MGGMAQDSPPQILVHPQDQLFQGPGPARMSCQASGQPPPTIRWLLNGQPLSMVPPDPHLLP
DGTLLLLQPPARGHAHDGQALSTDLGVYTCEASNRLGTAWSRGARLSAVLREDFQIQPRDM
VAVVGEOFTECGPPWGHPEPTVSWWDKGKPLALQPGHRTVSGSLLMARAEEKSDEGTYMCV
ATNSAGHRESRAARVSIQEPQDYTEPVELLAVRIQLENVTLLNPDPAEGPKPRPAVWLSWKV
SGPAAPAQSYTALFRQTAPGGQGAPWAEELLAGWQSAELGGLHWQDYEFKVRPSSGRARG
PDSNVLLLRLPEKVPAPPQEVTLKPGNGTVFVSWVPPPAENHNGIIRGYQVWSLGNTSLPP
ANWTVVGEQTQLEIATHMPGSYCVQVAVTGAGAGEEPSRPVCLLLEQAMERATQEPSEHGPW
TLEQLRATLKRPEVIATCGVALWLLLGTAVCIHRRRARVHLPGLYRYTSEDAILKHRMD
HSDSQWLADTWRSTSGSRDLSSSSLSSRLGADARDPLDCRRSLLWDSRSPGVPLLPDTST
FYGSLIAELPSSTPARPSPQVPAVRRLPPQLAQLSSPCSSSDSLCSRRGLSSPRLSIAPAEA
WKAKKKQELQHANSSPLLRGSHSLELRACELGNGRSKNLQS PGAVPQALVAWRALGPKLLS
SSNELVTRHLPPAPLFPHETPPTQSQQTQPPVAPQAPSSILLPAAPIPILSPCSPPSPQASS
LSGPSPASSRLSSSSLSSLGEDQDSVLTPEEVALCLELSEGEETPRNSVSPMPRAPSPTTY
GYISVPTASEFTDMGRTGGGVGPKGGVLLCPPRCLTPSEGSLANGWGSASEDNAASARA
SLVSSSDGSFLADAHFARALAVAVDSFGFGLEPREADCVFIDASSPPSPRDEIFLTPNLSLP
LWEWRPDWLEDMEVSHTQRLGRGMPPWPPDSQISSQLHCRMPKAGASPVDYS
```

Important features:

Transmembrane domain:

amino acids 448-467

N-glycosylation sites:

amino acids 224-227, 338-341, 367-370, 374-377, 658-661 and 926-929

N-myristylation sites.

amino acids 47-52, 80-85, 88-93, 99-104, 105-110, 181-186, 272-277, 290-295, 355-360, 403-408, 462-467, 561-566, 652-657, 849-854 and 876-881

Phosphotyrosine interaction domain proteins

amino acids 740-753

FIGURE 78

CTCCCACGGTGTCCAGGCCAGAATGCGGCTTCTGGTCTGCTATGGGTTGCCTGCTGCT
CCCAGGTTATGAAGCCCTGGAGGGCCCAGAGGAATCAGCGGGTCGAAGGGGACACTGTGT
CCCTGCAGTGCACCTACAGGGAAAGAGCTGAGGGACCACCGAAGTACTGGTGCAGGAAGGGT
GGGATCCTCTCTCGCTGCTCTGGCACCATCTATGCAGAAGAAGAAGGCCAGGAGACAAT
GAAGGGCAGGGTGTCCATCCGTGACAGGCCAGGAGCTCTGCTCATTGTGACCCGTGGA
ACCTCACCCCTGCAAGACGCTGGGAGTACTGGTGTGGGTCGAAAAACGGGCCCCGATGAG
TCTTACTGATCTCTGTTCTCGTCTTCCAGGACCCCTGCTGCTCCCTCCCTCTCCAC
CTTCCAGCCTCTGGCTACAACACGCCTGCAGGCCAAGGCAAAGCTCAGCAAACCCAGCCC
CAGGATTGACTCTCTGGCTCTACCCGGCAGCCACCAGCCAAGCAGGGAAAGACAGGG
GCTGAGGCCCTCCATTGCCAGGGACTTCCCAGTACGGGCACGAAAGGACTTCTCAGTACAC
AGGAACCTCTCCTCACCCAGCAGCACCTCCTCCTGCAGGGAGCTCCGCCCCCCATGCAGC
TGGACTCCACCTCAGCAGAGGACACCAGTCCAGCTCTCAGCAGTGGCAGCTCTAAGCCCAGG
GTGTCCATCCCAGTGGTCCGCATACTGGCCCCAGTCCTGGTGTGCTGAGCCTTCTGTCAGC
CGCAGGCCCTGATGCCTCTGCAGCCACCTGCTCCTGTGGAGAAAGGAAGCTAACAGGCCA
CGGAGACACAGAGGAACGAGAAGTCTGGCTCTCAGCAGTGGACTGCGGAGGAAAGGAAGCC
CCTTCCCAGGCCCTGAGGGGACGTGATCTCAGTGCCTCCACACATCTGAGGAGGA
GCTGGGCTCTCGAAGTTGTCTCAGCGTAGGGCAGGAGGCCCTCTGGCCAGGCCAGCAGT
GAAGCAGTATGGCTGGCTGGATCAGCAGGATTCCGAAAGCTTCCACCTCAGCCTCAGAG
TCCAGCTGCCCGACTCCAGGGCTCTCCCCACCCCTCCCCAGGCTCTCCTCTGCATGTTCA
GCCTGACCTAGAAGCGTTGTCAGCCCTGGAGGCCAGAGCGGTGGCTTGCTCTCCGGCTG
GAGACTGGACATCCCTGATAAGGTTACATCCCTGGCAGAGTACCAAGGCTGCTGACCTCA
GCAGGGCCAGACAAGGCTCAGTGGATCTGGTCTGAGTTCAATCTGCCAGGAACCTGGC
CTCATGCCAGTGTGGACCCCTGCCTCCTCCACTCCAGACCCCACCTTGTCTCCCTCCC
TGGCGTCTCAGACTTAGTCCCACGGTCTCCTGCATCAGCTGGTGTGAAGAGGAGCATGCT
GGGGTGAAGACTGGGATTCTGGCTTCTTTGAACCACCTGCATCCAGGCCCTCAGGAAGCCT
GTGAAAAACGTGATTCTGGCCCCACCAAGACCAACAAACCATCTGGGTTGGTGCAG
GACTCTGAATTCTAACAAATGCCAGTGAATGTCGCACTTGAGTTGAGGGCAGTGGCCTG
ATGAACGCTCACACCCCTTCAGCTTAGAGTCTGCATTGGGCTGTGACGTCTCCACCTGCC
CAATAGATCTGCTCTGTCTGCACACCAGATCCACGTGGGACTCCCTGAGGCCTGCTAAG
TCCAGGCCTGGTCAGGTGCACATTGCAGGATAAGCCCAGGACGGCACAGAAGTGG
TTGCCTTNCCATTGCCCTCCCTGGNCCATGCCTTGCCTTGAAAGGGGAAATGATGAAGA
AACCTTGGCTCTTGTCTGGAAAGGGTTACTTGCTATGGGTTCTGGTGGCTAGAGA
GAAAAGTAGAAAACCAGAGTGCACGTAGGTGTCAACACAGAGGAGTAGGAACAGGGCGG
ATACCTGAAGGTGACTCCGAGTCCAGCCCCCTGGAGAAGGGGTCGGGGTGGTAAAGTA
GCACAACACTATTTCATTTCATTATTATTGTGTTTAAGACAGAATCTCGTGCT
GCTGCCAGGCTGGAGTGCAGTGGCACGATCTGCAAACCTCCGCCTCTGGGTTCAAGTGATT
CTTCTGCCTCAGCCTCCGAGTAGCTGGGATTACAGGCACGCACCACACCTGGCTAATT
TTTGTACTTTAGTAGAGATGGGTTTACCATGTTGCCAGGCTGGTCTTGAACCTCTGAC
CTCAAATGAGCCTCTGCTTCAGTCTCCAAATTGCCGGATTACAGGCATGCCACTGTG
TCTGCCCTATTCTTTAAAAGTGAATTAAAGAGTTGTCAGTATGCAAAACTTGGAAAG
ATGGAGGAGAAAAGAAAAGGAAGAAAAATGTCACCCATAGTCTCACCAAGAGACTATCAT
TATTTCGTTTGTGACTTCCTCCACTCTTCTTCTTCACATAATTGCCGGTGTCTT
TTTACAGAGCAATTATCTGTATATAACACTTGTATCTGCTTACAGACCTTTATAAATAAA
ATGTTCATCA
GCTGCATAAAAAAAAAAAAAA

FIGURE 79

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44196
<subunit 1 of 1, 332 aa, 1 stop
<MW: 36143, pI: 5.89, NX(S/T): 1
MRLLVLLWGCLLPGYEALEGPEEISGFEGDTVSLQCTYREELRDHRKYWCRKGGILFSRCS
GTIYAAEEEGQETMKGRVSIRDSRQEELSLIVTLWNLTQDAGEYWCVEKRGPDESLLISLFV
FPGPCCPPSPSPTFQPLATTRLQPKAKAQQTQPPGLTSPGLYPAATTAKQGKTGAEAPPLPG
TSQYGHERTSQYTGTSPHPATSPAGSSRPPMQLDSTAEDTSPALSSGSSKPRVSIPMVR
LAPVLVLLSLLSAAGLIAFCSHLLWRKEAQQTETQRNEKFWLRLTAEEKEAPSQAPEGD
VISMPPLHTSEEELGFSKFVSA
```

Important features:

Signal peptide:

amino acids 1-17

Transmembrane domain:

amino acids 248-269

N-glycosylation site.

amino acids 96-99

Fibrinogen beta and gamma chains C-terminal domain.

amino acids 104-113

Ig like V-type domain:

amino acids 13-128

FIGURE 80

TTGTGACTAAAGCTGGCTAGCAGGCCAGGGAGTGCAGCTGCAGGCAGGGGTGGCAGGA
GCCGCAGAGCCAGAGCAGACAGCCGAGAACAGGTGGACAGTGTGAAAGAACAGTGGTCTC
GCTCTGTTGCCAGGCTAGAGTGTACTGGCGTGTACATAGCTCACTGCAGCCTCAGACTCCT
GGACTTGAGAAATCCTCCTGCCTAGCCTCCTGCATATCTGGACTCCAGGGGTGACTCAA
GCCCTGTTCTCTCCTCTGTGAGTGGACCACGGAGGCTGGTGAGCTGCCTGTACATCCAA
AGCTCAGCTCTGAGCCAGAGTGGTGGCTCCACCTCTGCCGCCGGCATAGAACGCCAGGAG
CAGGGCTCTCAGAAGGCAGGTGGTGCCTCAGCTGGGATCATGTTGTTGGCCCTGGTCTGTCTGC
TCAGCTGCCTGCTACCCCTCCAGTGAGGCCAAGCTCTACGGTCGTGTGAACTGGCCAGAGTG
CTACATGACTTCGGGCTGGACGGATACCGGGGATACAGCCTGGCTGACTGGGTCTGCCTTGC
TTATTCACAAGCGGTTCAACGCAGCTGCTTGACTACGAGGCTGATGGGAGCACCAACA
ACGGGATCTCCAGATCAACAGCCGGAGGTGGTGCAGCAACCTCACCCGAACGTCCCCAAC
GTGTGCCGGATGTACTGCTCAGATTGTTGAATCTAATCTCAAGGATACCGTTATCTGTGC
CATGAAGATAACCCAAGAGCCTCAGGGTCTGGTTACTGGGAGGCCTGGAGGCATCACTGCC
AGGGAAAAGACCTCACTGAATGGGTGGATGGCTGTGACTTCTAGGATGGACGGAACCATGCA
CAGCAGGCTGGAAATGTGGTTGGTCTGACCTAGGCTGGAAAGACAAGGCCAGCGAATA
AAGGATGGTTGAACGTGAAA

FIGURE 81

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52187
<subunit 1 of 1, 146 aa, 1 stop
<MW: 16430, pI: 5.05, NX(S/T): 1
MLLALVCLLSCLLPSSEAKLYGRCELARVLHDFGLDGYRGYSLADWVCLAYFTSGFNAAALD
YEADGSTNNGIFQINSRRWCSNLTPNVPNVCRMYSCLLNPNLKDVTICAMKITQEPMQGLGY
WEAWRHHCQGKDLTEWVDGCDF
```

Important features:

Signal peptide:

amino acids 1-18

N-myristoylation site.

amino acids 67-72

Homologous region to Alpha-lactalbumin / lysozyme C proteins.

amino acids 34-58 (catalytic domain), 111-132 and 66-107

FIGURE 82

AGCCGCTGCCCGGGCGCCCGCGGGCACCATGAGTCCCCGCTCGCCTGCGTT
GCTGCGCCTCCTCGTCTCGCCGTCTCTCAGCCGCGAGCAACTGGCTGTACCTGGCCA
AGCTGTCGTCGGTGGGGAGCATTCTCAGAGGAGGAGACGTGCGAGAAACTCAAGGGCTGATC
CAGAGGAGGTGCAGATGTGCAAGCGAACCTGGAAGTCATGGACTCGGTGCGCCGCGGTGC
CCAGCTGGCATTGAGGAGTGCCAGTACCAAGTCCGGAACCGGCGCTGGAACTGCTCACAC
TCGACTCCTGCCGTCTCGCAAGGTGGTGACGCAAGGGACTCGGAGGCAGCTCGT
TACGCCATCTCTCGGCAGGTGTGGCCTTGCAGTGACGCGGGCGTCAGCAGTGGGAGCT
GGAGAAGTGGGCTGTGACAGGACAGTCATGGGTGAGCCACAGGGCTTCCAGTGGTCAG
GATGCTCTGACAACATCGCCTACGGTGTGGCCTCTCACAGTCGTTGTGGATGTGCGGGAG
AGAACAGGGGGCCTGTCAGCAGAGCCCTCATGAACCTCCACAACAATGAGGCCGGCAG
GAAGGCCATCCTGACACACATGCGGGTGAATGCAAGTGCCACGGGTGTCAGGCTCCTGTG
AGGTAAAGACGTGCTGGCGAGCCGTGCCCTCCGCCAGGTGGTCACGCACTGAAGGAG
AAGTTGATGGTGCCACTGAGGTGGAGCCACGCCGTGGCTCCAGGGCACTGGTACCG
ACGCAACGCACAGTTCAAGCCGACACAGATGAGGACCTGGTACTTGGAGCCTAGCCCCG
ACTTCTGTGAGCAGGACATGCGCAGCGCGTGTGGCACGAGGGCCGCACATGCAACAAG
ACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGGCCGGCTTCCACACGGCGCA
GGTGGAGCTGGCTGAACGCTGCAAGCTGCAAATTCCACTGGTGTGCTTGTCAAGTGGCGC
AGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCATGACCGCCTGCCTAGCCCTGCCGGC
AACCACCTAGTGGCCCAGGGAGGCCATAATTAAACAGTCTCCACCACCTACCCAAGA
GATACTGGTTGTATTTTGTCTGGTTGGTTGGCTCATGTTATTATTGCCGAA
ACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGCTCCCCAAAGCCTGGGCTTGTGGCT
GCCACTGACCAAAGGGACCTGCTCGTGCCTGGCTGCCCATGTGGCTGCCACTGACCA
CTCAGTTGTTATCTGTGTCCGTTTCTACTTGCAACCTAAGGTGGAGTAACAAGGAGTAT
TACCACCATGGCTACTGACCGTGTCACTGGGAAGAGAGGGGCTTATGGCAGGGAAAATA
GGTACCGACTTGATGGAAGTCACACCCCTGGAAAAAAAGAACTCTAACTCTCCAGCACACA
TACACATGGACTCCTGGCAGCTTGAGCCTAGAACGCATGTCTCTCAAATGCCCTGAGAAAGG
GAACAAGCAGATAACCAGGTCAAGGGCACCAAGGGTCATTCAGCCCTACATGGACAGCTAGA
GGTCGATATCTGTGGTCCTCCAGGCAAGAAGAGGGAGATGAGAGCAAGAGACGACTGAA
GTCCCACCCCTAGAACCCAGCCTGCCCTGGGAAGAGGGAAACTTAACCACTCC
CCAGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCGGCTGTGCCCT
TGCAGTCATGCCCGAGTCACCTTCACAGCGCTGTTCCATGAAAATGAAAAACACACAC
AC
GAGAGGGAGGAAAGGGCTGTGCCCTTGCACTGCCAGTCACCTTCACAGCACTGTTCTC

FIGURE 83

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328
<subunit 1 of 1, 351 aa, 1 stop
<MW: 39052, pI: 8.97, NX(S/T): 2
MSPRSCLRSLLVFAVFSAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNL
VMDSVRRGAQLAIEECQYQFRNRRWCSTLDSLPGKVVVTQGTREAAFVYAISSAGVAFAV
TRACSSGELEKCGCDRTVHGVPQGFQWSGCSDNIAYGVAFSQSFVDVRERSKGASSRALK
NLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTCWRADVPPFRQVGHALKEKFDGATEVEPRR
VGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTSKAIDGCELLC
CGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR
```

Important features:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 88-91 and 297-300

Wnt-1 family signature.

amino acids 206-215

Homologous region to Wnt-1 family proteins

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

FIGURE 84

CGGACGCGTGGCGGACGCGTGGCGGACGCGTGGCGGACGCGTGGCTGGGTGCCTGCAT
CGCCATGGACACCAACCAGGTACAGCAAGTGGGGCGCAGCTCCGAGGAGGTCCCCGGAGGGC
CCTGGGGACGCTGGGTGCACTGGAGCAGGAGACCCCTCTTCTTGGCCCTGGCTGTCCGGTC
ACCACAGTCCTTGGGCTGTGATTCTGAGTATCCTATTGTCCAAGGCCTCACGGAGCGCAGC
GGCGCTGCTTGACGCCACGACCTGCTGAGGACAAACGCCCTCGAACAGCAGACGGCGCGCTGG
GTGCCCTGAAGGAGGAGGTCGGAGACTGCCACAGCTGCTGCCAGGGACGCAGGCGCAGCTG
CAGACCACGCGCGCGAGCTGGGAGGCCAGGGCGAGCTGATGGAGCAGGAGAGGCCCT
GCCGGAACTGCGTGAGCGCGTGACCCAGGGCTTGGCTGAAGCCGGCAGGGCCGTGAGGACG
TCCGCACTGAGCTGTTCCGGCGCTGGAGGCCGTGAGGCTCCAGAACAACTCCTGCGAGCCG
TGCCCCACGTCGTGGCTGTCCTCGAGGGCTCCTGCTACTTTCTCTGTGCCAAAGACGAC
GTGGCGGCAGGATCACTGCGCAGATGCCAGCGCAGCTGGTATCGTTGGGGCC
TGGATGAGCAGGGCTTCCTCACTCGAACACGCGTGGCCGTGGTACTGGCTGGGCCCTGAGG
GCTGTGCGCCATCTGGCAAGGTTAGGGCTACCAGTGGGTGGACGGAGTCTCTCAGCTT
CAGCCACTGGAACCAGGGAGAGCCCAATGACGCTGGGGCGCGAGAACTGTGTCATGATGC
TGCACACGGGCTGTGGAACGACGCACCGTGTGACAGCAGAAGGACGGCTGGATCTGTGAG
AAAAGGCACAACTGCTGACCCCGCCAGTGCCCTGGAGGCCGCCATTGCAGCATGCGTA
TCCTGGGGCTGCTCACCTCCCTGGCTCCTGGAGCTGATTGCCAAAGAGTTTTTCTTCCT
CATCCACCGCTGCTGAGTCTCAGAAACACTGGCCCAACATAGCCCTGTCCAGCCCAGTGC
TGGGCTCTGGACCTCCATGCCACCTCATCCTAACTCCACTCACGCAGACCCAACCTAAC
TCCACTAGCTCCAAAATCCCTGCTCCTGCGTCCCCGTGATATGCCTCCACTTCTCCCTAA
CCAAGGTTAGGTGACTGAGGACTGGAGCTGTTGGTTCTCGCATTTCACCAAACGGAA
AGCTGTTTGCAAGCCTGAGGAAGCATCAATAAAATTTGAGAAATGAAAAAA

FIGURE 85

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56352
<subunit 1 of 1, 293 aa, 1 stop
<MW: 32562, pI: 6.53, NX(S/T): 2
MDTTRYSKWGGSEEVPGGPWGRVHWSRRPLFLALAVLVTTVLWAVILSILLSKASTERAA
LLDGHDLLRTNASKQTAALGALKEEVGDCHSCCSGTQAQLQTTRAELGEAQAKLMEQESALR
ELRERVTQGLAEAGRGRREDVRTELFRALEAVRLQNNSCEPCPSTSWSFEGSCYFFSVPKTTW
AAAQDHCADASAHLVIVGGLDEQGFLTRNTRGRGYWLGLRAVRHLGKVQGYQWVDGVSLFS
HWNQGEPNDAWGRENVCVMMHLHTGLWNDAPCDSEKDGWICEKRHNC
```

Important features:

Type II transmembrane domain:

amino acids 31-54

N-glycosylation sites.

amino acids 73-76 and 159-162

Leucine zipper pattern.

amino acids 102-123

N-myristoylation sites.

amino acids 18-23, 133-138 and 242-247

C-type lectin domain signature.

amino acids 264-287

FIGURE 86

GCCAGGGAAAGAGGGTATCCGACCCGGGAAGGTCGCTGGCAGGGCGAGTTGGAAAGCG
GCAGCCCCCGCCGCCCGCAGCCCCTCTCCTCTTCTCCCACGTCTATCTGCCTCTCG
CTGGAGGCCAGGCCGTGCAGCATCGAAGACAGGAGGAACGGAGCTCATTGGCCGGCCGG
GGCGCCGGCTCGGGCTTAAATAGGAGCTCCGGCTCTGGCTGGGACCCGACCGCTGCCGGC
CGCGCTCCCGCTGCTCCTGCCGGTGATGGAAAACCCCAGCCCAGGCCGCCCCGGCAAG
GCCCTCTGCGCTCTCCTGCCACTCTCGGCCGCCGGCCAGCCTCTTGGGGAGAGTC
CATCTGTTCCGCCAGAGCCCCGGCAAATACAGCATCACCTCACGGCAAGTGGAGCCAGA
CGGCCTTCCCCAAGCAGTACCCCTGTTCCGCCCTGCGCAGTGGCTTCGCTGCTGGGG
GCCGCGCATAGCTCCGACTACAGCATGTGGAGGAAGAACAGTACGTCACTAACGGCTGCG
CGACTTGGGAGCGCGGGAGGGCTGGCGCTGATGAAGGAGATCGAGGCGGGAGGGAGG
CGCTGCAGAGCGTGCACGAGGTGTTCGCGCCCGCCGCCCCAGCGGCACCGGGCAGACG
TCGGCGAGCTGGAGGTGCAGCGCAGGCACTCGCTGGCTCGTTGTGGCGCATCGTGCC
CAGCCCCGACTGGTTCGTGGCGTGGACAGCCTGGACCTGTGCGACGGGACCGTTGGCGGG
AACAGGCGGCCTGGACCTGTACCCCTACGACGCCGGACGGACAGCGGCTCACCTCTCC
TCCCCCAACTCGCCACCATCCCGCAGGACACGGTGACCGAGATAACGTCTCCTCTCCAG
CCACCCGGCCAACTCCTTCTACTACCCGGCTGAAGGCCCTGCCTCCCATGCCAGGGTGA
CACTGCTGCCGTGCGACAGAGCCCCAGGGCTTCATCCCTCCCAGTCCTGCCAGC
AGGGACAATGAGATTGTAGACAGCGCCTCAGTTCCAGAAACGCCGCTGGACTGCGAGGTCTC
CCTGTGGTCGTCTGGGACTGTGCGGAGGCCACTGTGGAGGCTGGGACCAAGAGCAGGA
CTCGCTACGTCCGGTCCAGCCGAAACAACGGGAGCCCTGCCAGCTCGAAGAAAGAG
GCTGAGTGCCTGATAACTGCGTCTAAGACCAGAGCCCCGAGCCCTGGGCCCCCG
GAGCCATGGGTGTCGGGGCTCCTGTGAGGCTCATGCTGCAGCGGCCGAGGGCACAGGG
GGTTTGCCTGCTCCTGACCGCGGTGAGGCCGCCGACCATCTGCACTGAAGGCCCT
CTGGTGGCCGGCACGGCATTGGAAACAGCCTCCTCCTGCCAACCTTGCTTCTTGGGG
CCCCCGTGTCCCGTCTGCTCTCAGCCTCCTGCCAGGATAAAGTCATCCCCAAGGCTC
CAGCTACTCTAAATTATGTCTCCTTATAAGTTATTGCTGCTCCAGGAGATTGTCCTCATCG
TCCAGGGGCCTGGCTCCACGTGGTGCAGATAACCTCAGACCTGGTCTAGGCTGTGCTG
AGCCCACTCTCCCGAGGGCGCATCCAAGCGGGGCCACTTGAGAAGTGAATAAATGGGGCG
TTTGGGAAGCGTCAGTGTTCATGTTATGGATCTCTGCGTTGAATAAAGACTATCTCT
GTTGCTCACAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 87

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53971
><subunit 1 of 1, 331 aa, 1 stop
><MW: 35844, pI: 5.45, NX(S/T): 2
MENPSPAAALGKALCALLLATLGAAGQPLGGESICSARAPAKYSITFTGKWSQTAFFPKQYPL
FRPPAQWSSLGAAHSSDYSMWRKNQYVSNGLRDFAERGEAWALMKEIEAAGEALQSVHEVF
SAPAVPSGTGQTSAELEVQRRHSLVSVVRIVPSPDWFVGVDSDLCDGDRWREQAALDLYP
YDAGTDGFTFSSPNFATIPQDTVTEITSSSPSHPANSFYYPRLKALPPIARVTLLRLRQSP
RAFIPPPAPVLPSPRDNEIVDSASVPETPLDCEVSLWSSWGLCGGHCGRLGTSRTRYVRVQPA
NNGSPCPELEEEAECVPDNCV
```

Important features:

Signal peptide:

amino acids 1-26

FIGURE 88

GGCGGGCGTCCGTGAGGGGCTCCTTGGCAGGGTAGTGTGTTGGTGTCCCTGCTTGCCTGA
TATTGACAAACTGAAGCTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCCGA
CAGCTTAGTGGCCGGCCGCGCTCTCATCCCCGTAAGGAGCAGAGTCCTTGTACTGAC
CAAGATGAGCAACATCTACATCCAGGAGCCTCCCACGAATGGAAAGGTTTATTGAAAACTA
CAGCTGGAGATATTGACATAGAGTTGTGGTCAAAGAAGCTCCTAAAGCTTGAGAAATT
ATCCAACTTGTTGGAAGCTTATTGACAATACCATTTCATAGAGTTGTGCCTGGTT
CATAGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGAGAGTCTATCTATGGAGCGC
CATTCAAAGATGAATTTCATTACGGTTGCCTTAATCGGAGAGGACTGGTTGCCATGGCA
AATGCTGGTTCTCATGATAATGGCAGCCAGTTTCTTCACACTGGTCGAGCAGATGA
TAACAATAAGCATAACCCTTGGAAAGGTTACAGGGATACAGTATATAACATGTTGCGAC
TGTCAAGTAGACATTGATGATGACGAAAGACCAACATAATCCACACAAAATAAAAGCTGT
GAGGTTTGTAAATCCTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTAAAAAAGA
GAAACCAGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTAGTTACTTT
CATTGGAGAGGAAGCTGAGGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGA
GGCAAAAGCAAAAGTAGTCATGACTTGCTTAAGGATGATCCACATCTCAGTTCTGTTCCAGT
TGTAGAAAGTGAAGGAGGTGATGCACCAGATTAGTTGATGATGGAGAAGATGAAAGTGCAG
AGCATGATGAATATATTGATGGTGATGAAAAGAACCTGATGAGAGAGAAAGAATTGCCAAAAA
TTAAAAAAGGACACAAGTGCATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATC
AGTCAGCCGCAGTGAAGAGCTCAGAAAAGCAAGACAATTAAACGGAACTCTTAGCAG
CAAAACAAAAAAAGTAGAAATGCAGAAAACAAGCAGAAAAAAGAAGTGAAGAGGAAGAA
GCCCTCCAGATGGTGCTGTTGCCGAATACAGAAGAGAAAAGCAAAAGTATGAAGCTTGAG
GAAGCAACAGTCAAAGAAGGGAACTTCCGGAGATCAGACCCCTGCACTGCTGAACCA
TTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAAATGACATTCTGAAACAGAA
GTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTGAGGATAAAAGCAGAAAAGT
GAAAGATGCAAGCATGCAAGACTCAGATACATTGAAATCTATGATCCTCGGAATCCAGTGA
ATAAAAGAAGGAGGAAAGAAAGCAAAAGCTGATGAGAGAGAAAAAGAAAGAAGATAAAAT
GAGAATAATGATAACCAGAACCTGCTGGAAATGTGCCTACAATGCCCTGTAACAGCCATTG
TTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTGAAACCTGTTCTGGTTTG
AAAAACAATTATCTTGTGAAATTGTGGAATGATGTAAGCAAATGCTTTGGTTACTGG
TACATGTGTTTCTCTAGCTGACCTTATATTGCTAAATCTGAAATAAAACTTCC
TCCACAAAAAA

40017083 - 402404

FIGURE 89

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIYIQLQEPPTNGKVLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFI
VQGGDPTGTGSGGESIYGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFTLGRADELN
NKHTIFGKVTGDTVYNMLRLSEVDIDDDERPHNPKIKSCEVLFNPFDIIPREIKRLKKEK
PEEEVKKLKPKGTKNFSLLSFGEEAEEEEEVNRVSQSMKGKSCKSSHDLKDDPHLSSVPVV
ESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKLLKKDTSANVKSAGEGEVEKKSV
SRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEEAPPDGAVAAYRREKQKYEALRK
QOSKKGTSREDQTLALLNQFKSKLTQAIATPENPDIPETEVEDDEGWMSHVLQFEDKSRKVK
DASMQDSDTFFEIYDPRNPVNKRREESKKLMREKKERR
```

Important features:

Signal peptide:

amino acids 1-21

N-glycosylation sites.

amino acids 109-112 and 201-204

Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.

amino acids 49-66

Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase

amino acids 96-140, 49-89 and 22-51

FIGURE 90

CGCCGCCGTTGGGGCTTGGAAAGTTCCCGCCAGGTCCGTGCCGGCGAGAGAGATGCTGCCCGG
CCCGCCTCGGCTTGAGGCGAGAGAAGTGTCCCAGACCCATTGCGCTTGACGGCGTCG
AGCCCTGGCCAGACATGTCCACAGGGTTCTCCTCGGGCCGGGACTCTGGGCTCCACCACC
GTGGCCGCCGGCGGGACCAAGCAGGGCGTTTCTCCTCGGAACGGGAACGTCTAGCAA
CCCTCTGTGGGGCTCAATTGGAAATCTTGGAAAGTACTTCAACTCCAGCAACTACATCTG
CTCCTTCAAGTGGTTTGGAACCGGGCTTTGGATCTAAACCTGCCACTGGGTTACTCTA
GGAGGAACAAATAACAGGTGCCTGCACACCAAGAGGCCTCAAGTGGTACCAAATATGGAAC
CCTGCAAGGAAAACAGATGCATGTGGGAAGACACCCATCCAAGTCTTTAGGAGTCCCCT
TCTCCAGACCTCCTCTAGGTATCCTCAGGTTGCACCTCCAGAACCCCCGGAGCCCTGGAAA
GGAATCAGAGATGCTACCACCTACCCGCTGGATGGAGTCTCGCTCTGTCGCCAGGCTGGAG
TGCAGTGGCACGATCTCGGCTCACTGCAACCTCCGCTCCGGTTCAAGCGAGTCTCCCTGC
CTCAGCCTCTGAGTGTCTGGGCTACAGGTGCCTGCAGGAGTCTGGGCCAGCTGGCCTCG
ATGTACGTCAAGCACGCCGGAACGGTACAAGTGGCTGCCTCAGCGAGGACTGTCTGTACCT
GAACGTGTACGCCGGCGCGCGCCGGGATCCCCAGCTGCCAGTGTATGGTCTGGTTCC
CGGGAGGCGCCTTCATCGTGGCGCTGCTTCTCGTACGAGGGCTCTGACTTGGCGCCCGC
GAGAAAGTGGTGTGGTGTTCAGCAGCACAGGCTGGCATCTTGGCTTCTGAGCACGGA
CGACAGCCACGCCGGCGCGGGAACTGGGGCTGCTGGACCAGATGGGGCTCTGCGCTGGGTGC
AGGAGAACATCGCAGCCTCGGGGGAGACCCAGGAAATGTGACCCTGTTGGCCAGTCGGCG
GGGCCATGAGCATCTCAGGACTGATGATGTCACCCCTAGCCTGGGTCTTCCATCGGGC
CATTCCCAGAGTGGCACCGCGTTATTCAAGACTTTCATCACTAGTAACCCACTGAAAGTGG
CCAAGAAGGTTGCCACCTGGCTGGATGCAACCACACAGCACAGATCCTGGTAAACTGC
CTGAGGGCACTATCAGGGACCAAGGTGATGCGTGTGTCACAGGAGATTCTCTCCA
GAACCTCCAGAGAGACCCGGAAGAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGG
TGATCCCAGATGACCCCTTGGTGTCTGACCCAGGGGAAGGTTCATCTGTGCCCTACCTT
CTAGGTGTCAACAACCTGGAATTCAATTGGCTCTGCCTATAATATCACCAAGGAGCAGGT
ACCACTTGTGGAGGAGTACCTGGACAATGTCAATGAGCATGACTGGAAGATGCTACGAA
ACCGTATGATGGACATAGTTCAAGATGCCACTTCGTGTATGCCACACTGCA
TACCAACGAGAAACCCAAATGATGGGAATCTGCCCTGCTGGCCACGCTACAACAAGGATGAA
AAGTACCTGCAGCTGGATTACCAAGAGTGGCATGAAGCTCAAGGAGAAGAAGATGGC
TTTTGGATGAGTCTGTACAGCTCAAAGACCTGAGAAGCAGAGGCAATTCTAAGGGTGGC
TATGCAGGAAGGAGCCAAAGAGGGTTGGCCCCACCATCCAGGCCCTGGGAGACTAGCCA
TGGACATACCTGGGACAAGAGTTCTACCCACCCAGTTAGAACTGCAGGGACTCCCTGCT
GCCTCCAGGCCAAAGCTAGAGCTTGTGCTGTTGTGGACCTGCACTGCCCTTCCAGCC
TGACATCCCAGTGTGCCCCCTACTTCAGTGTGACATCCAGTTAGGCCAGGCCCTGTCAAC
ACCACACTGTGCTCAGCTCCAGCAGAACCTCTTTTCCCTTCTCAAATCCT
CCCACCCCTCAATGTCTCCTGTGACTCCTCTTATGGGAGGTGACCCAGACTGCCACTGC
CCCTGTCAGTGCACCCAGCTGGCATTACCATCCACTGCTCAACCTGTTCTGTCTGT
TCACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACATGGAGCAAACCTTGGTAGTTGGGA
TCTCTCTCCCACCCACACTTATCTCCCCAGGGCCACTCCAAAGTCTATAACACAGGGTGG
TCTCTCAATAAAGAAGTGTGATTAGAAAAAAAAAAAAA

FIGURE 91

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATT SAPSSG
FGTGLFGSKPATGFTLGGTNTGALHTKRPQVVTKYGTLQGKQMHVGKTPIQVFLGVFSRPP
LGILRFAPPEPPEPWKGIRDATTYPPGWSLALSPGWSAVARSRLTATSASRVQASLLPQPLS
VWGYRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAF
IVGAASSYEGSDLAAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIA
AFGGDPGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVA
HLAGCNHNSTQILVNCLRALSGTKVMRVSNKMRFLQLNFQRDPEIIIWMSSPVVDGVVIPDD
PLVLLTQGVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEYLDNVNEHDWKMRLRNRMMD
IVQDATFVYATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA

Important features:

Signal peptide:

amino acids 1-29

Carboxylesterases type-B serine active site.

amino acids 312-327

Carboxylesterases type-B signature 2.

amino acids 218-228

N-glycosylation sites.

amino acids 318-321, 380-383 and 465-468

FIGURE 92

GAGAACAGGCCCTGCTCAGGCAGGCCCTGCGCCTCCTATGCGGAGATGCTACTGCCACTGCT
GCTGTCCTCGCTGCTGGCGGGTCCCAGGCTATGGATGGGAGATTCTGGATACGAGTGCAGG
AGTCAGTGTGGTGCAGGGCCTGTGCATCTCTGTGCCCTGCTTTCTCCCTACCCCCGA
CAAGACTGGACAGGGTCTACCCCAGCTATGGCTACTGGTTCAAAGCAGTGACTGAGACAAC
CAAGGGTGCTCCTGTGCCACAAACCACAGAGTCAGAGGGTGGAAATGAGCACCCGGGGCC
GATTCCAGCTCACTGGGATCCCAGCAAGGGAACTGCTCCTGGTATCAGCAGTGCACAGCAG
ATGCAGGATGAGTCACAGTACTTCTTCGGGTGGAGAGAGAGAAGCTATGTGACATATAATTT
CATGAACGATGGGTTCTTCTAAAAGTAACAGTGCTCAGCTCACGCCAGACCCCCAGGACC
ACAACACCGACCTCACCTGCCATGTGGACTTCTCCAGAAAGGGTGTGAGCGCACAGAGGACC
GTCCGACTCCGTGTGGCCTATGCCAGAGACCTGTTATCAGCATTACGTGACAACAC
GCCAGCCCTGGAGCCCCAGGCCAGGGAAATGTCCCACCTGGAAGCCCAAAAGGCCAGT
TCCTGCGGCTCCTCTGTGCTGACAGCCAGCCCCCTGCCACACTGAGCTGGTCTGCAG
AACAGAGTCCCTCTCCCGTCCATCCCTGGGCCCTAGACCCCTGGGCTGGAGCTGCCCG
GGTGAAGGCTGGGATTCAAGGCGCTACACCTGCCAGCGAGAACAGGCTGGCTCCAGC
AGCGAGCCCTGGACCTCTGTGCAGTATCCTCAGAGAACCTGAGAGTGATGGTTCCAA
GCAAACAGGACAGTCCCTGGAAAACCTTGGAACGGCACGCTCTCCAGTACTGGAGGGCA
AAGCCTGTGCCCTGGTCTGTGTCACACACAGCAGCCCCCAGCCAGGCTGAGCTGGACCCAGA
GGGACAGGTTCTGAGCCCCCTCCAGGCCCTCAGACCCGGGTCCTGGAGCTGCCCTGGGTT
CAAGTGGAGCACGAAGGAGAGTTCACCTGCCACGCTGGCACCCACTGGCTCCAGCACGT
CTCTCTCAGCCTCTCCGTGCACTATAAGAAGGGACTCATCTCAACGGCATTCTCCAACGGAG
CGTTCTGGGAATCGGCATCACGGCTCTCTTCTGCTGCCCTGATCATCATGAAG
ATTCTACCGAAGAGACGGACTCAGACAGAAACCCGAGGGCCAGGTTCTCCGGCACAGCAC
GATCCTGGATTACATCAATGTGGTCCCGACGGCTGGCCCCCTGGCTCAGAAGCGGAATCAGA
AAGCCACACCAAAACAGTCCCTGGACCCCTCCACCAGGTGCTCCCTCCCCAGAATCAAAG
AAGAACAGAAAAAGCAGTATCAGTTGCCAGTTCCCAGAACCCAAATCATCCACTCAAGC
CCCAGAATCCCAGGAGAGCCAAGAGGAGCTCATTATGCCACGCTCAACTTCCAGGCGTCA
GACCCAGGCCCTGAGGCCGGATGCCAAGGGCACCCAGGCGGATTATGCAGAAGTCAAGTC
CAATGAGGGTCTCTAGGCTTAGGACTGGACTTCGGCTAGGGAGGAAGGTAGAGTAAGAG
GTTGAAGATAACAGAGTGCAGGAAAGTTCTTCTCTCCCTCTCTCTCTCTCTCTCTCT
CTCTCTTCTCTCTCTTAAAAAAACATCTGGCCAGGGCACAGTGGCTCACGCCTGTAATC
CCAGCACTTGGGAGGGTTGAGGTGGGAGATCGCCTGAGGTGGAGTTGAGACCCAGCCTG
GCCAACTTGGTGAACCCCCGTCCTACTAAAAATACAAAAATTAGCTGGGATGGTGGCAGG
CGCCGTAAATCCTACCTACTTGGGAAGCTGAGGCAGGAGAACATCACCTGAACCTGGGAGACGG
AGGTTGCAGTGAGCCAAGATCACACCATTGCACGCCAGCCTGGCAACAAAGCAGACTCCA
TCTCAAAAAAAATCCTCAAATGGGTTGGGTGTCTGTAATCCCAGCACTTGGGAGGCTA
AGGTGGGTGGATTGCTTGAGCCCAGGAGTTGAGACCAGCCTGGCAACATGGTAAACCC
ATCTCTACAAAAATACAAAACATAGCTGGCTTGGTGGTGTGCTGTAGTCCCAGCTGT
CAGACATTAAACCAGAGCAACTCCATCTGAATAGGAGCTGAATAAAATGAGGCTGAGACC
TACTGGGCTGCATTCTCAGACAGTGGAGGCATTCTAAGTCACAGGATGAGACAGGAGGTCCG
TACAAGATAACAGGTATAAAAGACTTGTGATAAAACAGATTGCAAGTAAAGAAGCCAACCAA
ATCCCACCAAAACCAAGTTGGCCACGAGAGTGACCTCTGGTCGTCTCACTGCTACACTCCT
GACAGCACCATGACAGTTACAAATGCCATGGCAACATCAGGAAGTTACCCGATATGTCCCA
AAAGGGGGAGGAATGAATAATCCACCCCTTGTGTTAGCAATAAGCAAGAAATAACCATAAAA
GTGGGCAACCAGCAGCTCTAGGCGCTGCTTTGTCTATGGAGTAGGCCATTCTTGTGCTT
TACTTCTTAATAAAACTTGTCTTCACCTTAAAAAA

FIGURE 93

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1 stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSYPRQDWTGSTPAYGYWFK
AVTETTKGAPVATNHQSREVEMSTRGRFQLTGDPAKGNCISLIRDAQMDESQYFFRVERGS
YVTYNFMNDGFFLKVTVLSFTPQDHNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVIS
ISRDNTPALEPQPQGNVPYLEAQKGQFLRLCAAQSPATLSWVLQNRVLSSHPWGPRPL
GLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPHENRVMVSQANRTVLENLNGNTSL
PVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHP
LGSQHVSLSLSVHYKKGLISTAFSNGAFLGIGITALFLCLALIIMKILPKRRTQTEPRPR
FSRHSTILDYINVVPTAGPLAQKRNQKATPNSPRTPPPGAPSPESKKNQKKQYQLPSFPEP
KSSTQAPESQESQEELHYATLNFPGVRPRPEARMPKGTQADYAEVKFQ
```

Important features:

Signal peptide:

amino acids 1-15

Transmembrane domain:

amino acids 399-418

N-glycosylation site.

amino acids 100-103, 297-300 and 306-309

Immunoglobulins and major histocompatibility complex proteins signature.

amino acids 365-371

FIGURE 94

TGAAGAGTAATAGITGGAATCAAAAGAGTCAACGCAATGAACTGTATTTACTGCTCGT TT
TATGTTGGGAATTCCCTCTCCTATGGCCTGTCTGGAGCAACAGAAAACTCTCAAACAAAGA
AAGTCAAGCAGCCAGTGCATCTCATTGAGAGTGAAGCGTGGCTGGGTGGAACCAATTT
TTTGTACCAGAGGAAATGAATACGACTAGTCATCACATCGGCCAGCTAAGATCTGATTAGA
CAATGGAAACAATTCTTCCAGTACAAGCTTTGGGAGCTGGAGCTGGAAAGTACTTTATCA
TTGATGAAAGAACAGGTGACATATGCCATACAGAAGCTGATAGAGAGGAGCGATCCCTC
TACATCTTAAGAGCCCAGGTAATAGACATCGCTACTGGAAGGGCTGTGGAACCTGAGTCTGA
GTTTGTCAAAAGTTGGATATCAATGACAATGAACCAAATTCTAGATGAACCTTATG
AGGCCATTGTACCAGAGATGTCCAGAAGGAACATTAGTTATCCAGGTGACAGCAAGTGAT
GCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTACAGCTTACTTCAGGCCAGGCC
ATATTTCTGTTGAACCAACAAACAGGAGTCATAAGAATATCTCTAAAATGGATAGAGAAC
TGCAAGATGAGTATTGGGTAAATCATTCAAGCAAGGACATGATTGGTCAGCCAGGAGCGTTG
TCTGGAACAAACAAGTGTATTAATTAAACTTCAGATGTTAATGACAATAAGCCTATATTAA
AGAAAGTTATACCGCTTGACTGTCTGAAATCTGACCCACTGGACTTCTATAGGAACAA
TCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTACAGCATTGAAGAGGAT
GATTGCAAACATTGACATTACTAATCATGAAACTCAAGAAGGAATAGTTATTTAAA
AAAGAAAGTGGATTTGAGCACCAGAACCAACTACGGTATTAGAGCAAAAGTTAAAACCATC
ATGTTCTGAGCAGCTCATGAAGTACCAACTGAGGCTTCCACCACCTTCAAGATCCAG
GTGGAAGATGTTGATGAGCCTCTCTTCCCTCCATATTATGTTGAAGTGGATTTGA
AGAAACCCACAGGGATCATTGTTAGGCGTGGTCTGCCACAGACCCAGACAATAGGAAT
CTCCTATCAGGTATTCTATTACTAGGAGCAAAGTGTCAATATCAATGATAATGGTACAATC
ACTACAAGTAACTCACTGGATCGTGAATCAGTGCTGGTACAACCTAAGTATTACAGCCAC
AGAAAAATACAATATAGAACAGATCTTCGATCCCAGTGTATGCAAGTTCTAACATCA
ATGATCATGCTCCTGAGTTCTCAATACTATGAGACTTATGTTGTGAAATGCAGGCTCT
GGTCAGGTAAATTCAAGACTATCAGTGCACTGGATAGAGATGAATCCATAGAACGACCA
TTACTTTAATCTATCTGAGAACACTAACAAATTCAAGTTTACAATCATAGATAATCAAG
ATAACACAGCTGTCAATTGACTAATAGAACACTGGTTAACCTCAAGAACCTGCTTC
TACATCTCCATCTTAATTGCCGACAATGGAATCCCCTCAACTAACAGAACCCCTTAC
CATCCATGTCGTGACTGTGGTACAGTGGAGCACACAGACCTGCCAGTACCAAGGAGCTTG
TGCTTCCATGGGATTCAAGACAGAACAGTTATCATTGCTATTCTCATTGCAATTGATCATA
TTTGGGTTATTTTTGACTTTGGTTAAAACAACGGAGAAAACAGATTCTATTCTGA
GAAAAGTGAAGATTTAGAGAGAATATATTCAATATGATGATGAAGGGGGTGGAGAAGAAC
ATACAGAGGCCTTGATATAGCAGAGCTGAGGAGTAGTACCAATGCGGGAACGCAAGACT
CGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTGCAAGTTGGCCCCGA
CAGGCCATATTCAAGGAAATTCTTCTGAAAGCTCGAAGAACGCTAAACTGATCCGTGTG
CCCCCTTCTGATTCCCTCCAGACCTACGCTTTGAGGGAACAGGGTCATTAGCTGGATCC
CTGAGCTCCTAGAATCAGCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTT
GGGACCTCGCTTAAAAGATTAGCATGCATGTTGGTCTGCAGTGCAGTCAAATAATTAGG
GCTTTTACCATCAAATTAAAAGTGTAAATGTGTATTGCAACCCAATGGTAGTCTTAA
AGAGTTTGTGCCCTGGCTCATGGCGGGAAAGCCCTAGTCTATGGAGTTCTGATTCC
CTGGAGTAAATACTCCATGGTTATTTAAGCTACCTACATGCTGTCAATTGAAACAGAGATGTG
GGGAGAAATGAAACAATCAGCTCACAGGCATCAATACAACCAAGATTGAAGTAAAATAATG
TAGGAAGATATTAAAAGTAGATGAGAGGACACAAGATGTAGTCGATCCTATGCGATTATAT
CATTATTACTTAGGAAAGAGTAAAATACCAACGAGAAAATTAAAGGAGAAAATTG
CAAGTCAAATAGAAATGTACAAATCGAGATAACATTACATTCTATCATATTGACATGAAA
ATTGAAAATGTAGTCAGAGAAATTTCATGAATTATTCCATGAAGTATTGTTCTTAT
TTAAA

FIGURE 95

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGWVWNQFFVPEEMNTSHH
IGQLRSDLNDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSDINDNEPKFLDEPYEAIVPEMSPEGLVQVTASDADDPSSGNARL
LYSLLQGQPYFSVEPTTGVRISSKMDRELQDEYWVIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPPIFKESLYRLTVSESAPGTGSIGTIMAYNDIGENAEMDSIEEDDSQTFDIITNHE
TQEGLIVLKKVDFEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYVFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFNINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEHHFYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDCGDGSQTQYQELVLSMGFKTEVIIAILICIMIIFGFIFLTGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIRKFILEKLEEANTDPCAPPFDLQTYAFEGTGSLAGSLSSLESAVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518,
516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 96

ATTTCAGGCCAGCCATATTTNTGTTAACCAACAACAGGAGTCATAAGAATATTTNTA
AAATGGATAGAGAACTGCAAGATGAGTATTGGTAATCATTCAAGCCAAGGACATGATTGGT
CAGCCAGGAGCGTTGNTGGAACAACAAGTGTATTAATTAAACTTCAGATGTTAATGACAA
TAAGCCTATATTTAAAGAAAGTTATACCGCTTGACTGTTNTGAATCTGCACCCACTGGGA
NTTNTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAATGCAGAAATGGATTAC
AGCATTGAAGAGGGATGATTGCAAACATTGACATTATT

FIGURE 97

GCAACCTCAGCTCTAGTATCCAGACTCCAGGCCGCCCCGGCGGGACCCAAACCCGAC
CCAGAGCTTCTCCAGCGGGCGCAGCGAGCAGGGCTCCCGCCTTAACCTCCCTCCGGGG
CCCAGCCACCTTCGGGAGTCCGGGTTGCCACCTGCAAACCTCTCCGCCTCTGCACCTGCCA
CCCCTGAGCCAGCGCGGGCCCCGAGCGAGTCATGGCCAACGCGGGGCTGCAGCTGTTGGC
TTCATTCTCGCCTCTGGGATGGATCGGCGCATCGTCAGCACTGCCCTGCCCATGGAG
GATTACTCCTATGCCGGCACAACATCGTACCGCCAGGCCATGTACGAGGGGCTGTGGA
TGTCCCTGCGTGTGCAAGAGCACCGGGAGATCCAGTGCAAAGTCTTGACTCCTGCTGAAT
CTGAGCAGCACATTGCAAGCAACCGTGCCTTGATGGTGGCATCCTCTGGGAGTGAT
AGCAATCTTGTGGCCACCGTTGGCATGAAGTGATGAAGTGCTTGGAAAGACGATGAGGTGC
AGAAGATGAGGATGGCTGTCATTGGGGTGCATATTCTCTTGCAAGGTCTGGCTATTGTA
GTTGCCACAGCATGGTATGGCAATAGAATCGTCAAGAATTCTATGACCCATGACCCAGT
CAATGCCAGGTACGAATTGGTCAGGCTCTTCACTGGCTGGCTGCTCTCTGCC
TTCTGGGAGGTGCCCTACTTTGCTGTTCTGCCCCAAAACACCTCTTACCAACACCA
AGGCCCTATCCAAAACCTGCACCTCCAGCGGGAAAGACTACGTGTGACACAGAGGCAAAG
GAGAAAATCATGTTGAAACAAACGAAAATGGACATTGAGATACTATCATTAAACATTAGGAC
CTTAGAATTGGGTATTGTAATCTGAAGTATGGTATTACAAAACAAACAAACAAAAAA
ACCCATGTGTTAAAATACTCAGTGTAAACATGGCTTAATCTTATTGTTATCTCTTCTCA
ATATAGGAGGGAAAGATTTCCATTGTATTACTGCTTCCATTGAGTAATCATACTCAAAT
GGGGGAAGGGGTGCTCTTAAATATATAGATATGTATATACATGTTTCTATTAAAAA
ATAGACAGTAAAATACTATTCTCATTATGTTGATACTAGCATACTAAAATATCTCTAAAAT
AGGTAAATGTTAATTCCATTGATGAAGATGTTATTGGTATATTCTTTCTGTCC
TTATATACATATGTAACAGTCAAATATCATTACTCTTCTTCAATTAGCTTGGGTGCCCTTG
CCACAAGACCTAGCCTAATTACCAAGGATGAATTCTTCATTCTCATGCGTGCCCTTT
CATATACTTATTGTTTACCATTAATCTTATAGCATTGCACTGCTTATTAGCCCTTAT
TTGTTTGTGTTCAATTGGTCTCTATCTCCTGAATCTAACACATTCTAGCCTACATTAA
GTTCTAAAGCCAAGAAGAATTATTACAAATCAGAACTTGGAGGCAAATCTTCTGCATG
ACCAAAGTGATAAAATTCTGTTGACCTTCCCACACAATCCCTGACTCTGACCCATAGCACT
CTTGTGTTGCTTGAATAATTGTCAATTGAGTAGCTGCTGTTCCCCAGGTGTTG
AACACAACTTATTGATTGAATTTTAAGCTACTTATTGTTATTGTTCCAAAGTGTAAATTATCATGCGTTTA
TATCTCCTAATAAGGTGTGGTCTGTTGTGAACAAAGTGTAGACTTCTGGAGGTGATA
ATCTGGTACAATATTCTCTGTAGCTGTAAAGCAAGTCACCTAATCTTCTACCTCTTT
TTCTATCTGCCAAATTGAGATAATGATACTTAACCAGTTAGAAGAGGTAGTGTGAATATTAA
TTAGTTTATATTACTCTTATTCTTGAAACATGAACTATGCTATGTAGTGTCTTTATTGCT
CAGCTGGCTGAGACACTGAAGAAGTCAGTGAACAAAACCTACACACGTACCTCATGTGATT
CACTGCCCTCCTCTCTTACCACTGAACAAAACCTACACACATACCTTCA
GTGGTTAGTGCCTCCTCTCTTACCACTGAACAAAACCTACACACATAC
CTTCATGTGGCTCAGTGCCTCCTCTTACCACTGAACAAAACCTACGACACATAC
GACATGTTGTGCTCTGTTCCATTAAACAACGTCTTACTTTCCAGTCAGTACAGAATG
CTATTCACTTGAGCAAGATGATGTAATGAAAGGGTGTGGCAGTGGTGTCTGGAGACCTG
GATTGAGTCTGGGTGCTATCAATCACCCTGCTGTGTTGAGCAAGGCATTGGCTGCTGTA
GCTTATTGCTTCATCTGTAAGCGGTGGTTGTAATTCTGATCTTCCCACCTCACAGTGATG
TTGTGGGATCCAGTGAGATAGAATACTGTAAGTGTTGTTGTAATTAAAAAGTGCTAT
ACTAAGGGAAAGAATTGAGGAATTACTGCATACGTTTGGTGTGCTTCAAATGTTGA
AAATAAAAAAAATGTTAAG

FIGURE 98

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185
><subunit 1 of 1, 211 aa, 1 stop
><MW: 22744, PI: 8.51, NX(S/T): 1
MANAGLQLLGFI LAFLGWIGAIVSTALPQWRIYSYAGDNIVTAQAMY EGLWMSCVSQSTGQI
QCKVFDSSLNLSSTLQATRALMVVGILLGVIAIFVATVGMKCMKLEDDEVQKMRMAVIGGA
IFLLAGLAILVATAWYGNRIVQE FYDPMTPVNARYEFGQALFTGWAASLCLLGGALLCCSC
PRKTTSYPTPRPYPKPAPSSGKDYV
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domains:

amino acids 82-102, 118-142 and 161-187

N-glycosylation site.

amino acids 72-75

PMP-22 / EMP / MP20 family proteins

amino acids 70-111

ABC-2 type transport system integral membrane protein

amino acids 119-133

FIGURE 99

TTCTGGCCAAACCCGGGGCTNCAGCTGTTGGCTTCATCTGCCCTCCTGGGATGGATCGC
GCCATCNTCACACTGCCCTCCCCAGTGGAGGATTTACTCCCTATGCTGGCGACAAACATCG
TGACCGCCCAGCCCATGTACGAGGGCTGTGGATGTCNGCGTGTGCAGAGCACCAGGGCAG
ATCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCCGTGC
CTTGATGGTGGTTGGCATCCTCCTGGAGTGATAGCAATCTTGTGCCACCGTTGGCATGA
AGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGC
GCGATATTCTTCTGCAGGTCTGGCTATTTAGTTGCCACAGCATGGTATGCCAATAGAAN
CNTTCAACANTCTATGACCCATGACCCAGTCAATGCCAGGTACGAATTGGTCA
GGCTCTTCACTGGCTGGCTGCTGCTCTCTGCCCTCTGGGAGGTGCCCTACTTGCT
GTTCCCTGTCCC

FIGURE 100

ACCCCTGACCCAACGCGGCCCCCGACCGNTTCATGGCAAACGCGGGNCTCCAGCTGTTGG
GCTTCATTCTCCCCTTCCTGGGATGGACCGGCGCCATCNTCAGCACTGCCCTGCCAGTG
GAGGATTACTCCTATNCCGGCNACAAACATCGTACCGCCAGGCCNTGTACGAGGGGCTGT
GGATGTCCTGCGTGTGCAGAGCACCAGGAGATCCAGTGCAAAGTCTTGACTCCCTGCT
GAATCTGAGCAGCACATTGCAAGCAACCGTGCCTGATGGTGGTGGCATCCTCCTGGAG
TGATAGCAATCTNNNTGCCACCGTTGTNNNTGAAGTGTATGAAGTGCTTGGAAAGACGATGA
GGTGCAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATTTCTTGCAGGTCTGGCTA
TTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTATGACCCATGACCGA

FIGURE 101

GGGCCGACCATTATCCAACCGGGNTCACTGTTGGCTCATCTCCCTCCTGGATGAANCAGCGC
CATCNCAGACTCCCTGCCCATGGAGATTTNNCCTATGCTGGCGACAACATCNTGACCCCC
AGCCATGTACGAGGGGCTTGAACGTCNGCGTGTGCGAGANCACCGGGCAGATCCAGTGCAA
AGTCTTGACTCCTGCTGAATCTGNGCAGCACATTGCAGCAACCCNTGCCCTGATGGTGGT
TGGCATCCTCCTGGGAGTGATAGCAATCTTGTGGCCACCGTTGGCATGAAGTGTATGAAGT
GCTTGGAAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCATTGGGGCGCGATATTCTT
CTTGCAGGTCTGGCTATTNNNNGTGCACAGCATGGTATGGCAATAGAATCGTTCAAGAAT
TCTATGACCCCTATGACCCCAAGTCAATGCCAGGTACGAATTGGTCAGGCTCTCTCACTGGC
TGGGCTGCTGCTTCTCTGCCTCTGGGAGGTGCCCTACTTGCTGTTCTGCAGA

FIGURE 102

ATTCTCCCTCCTGGATGGATCGCNCCACCGTCACATTGCCTTCCCCANTGGAGGATTNAC
TCCTATGCTGGCGACAACATCGTACCCCCCAGGCCATTACCGAGGGCTTGGATGTCNT
GCNTGTCGAGAGCACCGGGCAGATCCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAG
CAGCACATTGCAAGCAACCGTGCCTGATGGGTTGGCATCCTCCTGGAGTGATAGCAAC
CTTGTCGCCACCGTTGGCATGAAGTGTATGAAGTGCTTGAAGACGATGAGGTGCCAGAAG
ATGAGGATGGCTGTCATTGGGGCGCGATATTCTTGCAGGTCTGGCTATTTAGTNGC
CACAGCATGGTATGGCAATAGANTNNTCNNNGNNNTCTATGACCCCTATGACCCAGTCAATG
CCAGGTACGAATTGGTCAGGCTCTCTCACTGGCTGGCTGCTGCTCTGCCTTCTG
GGAGGTGCCCTACTTGCTGTTCTGTCCC

FIGURE 103

AGAGCACCGGCAGATCCCAGTNCAAAGTCTTGACCCCTGCTGAATCTGAGCAGCACATTNC
AAGCAACCCCTTGCCTTGAAGGTGGTTGNCATCCCCCTGGGAGTGAATAGCAATCTTGTG
GCCACCCTGGCATGAAGTNTATGAAGTGCTTGAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTGCAGGTCTGGCTATTTAGTNNCCACAGCAT
GGTATGGCAATAGNATNNTCGNGGNTCTATGACCCATGACCCAGTCAATGCCAGGTAC
GAATTGGTCAGGCTCTCTTCACTGGCTGGCTGCTGCTCTCTGCCTCTGGGAGGTGC
CCTACTTGTGTTCTGTCCCCGAA

FIGURE 104

AGCAATGCCCTGCCCCCAGTGGAGGATTAATTCTATGNTGGGACAACATTGTGACNGCCC
AGGCCATGTACGGGGGCTGTGGATGTCCTGCGTGTGCGAGCACCAGGAGATCCAGTGC
AAAGTNTTGACTCCTGCTGAATTGAGCAGCACATTGCAAGCAACCCGTGCCTTGATGGT
GGTGGCATCTCCTGGAGTGATAGCAATCTTGTGCCACCGTGGNAATGAAGTGTATGA
AGTGCTTGGAAAGACGATGAGGTGCGAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATT
CTTNTTGCAAGGTCTGGCTATTTAGTTGCCACAGCATGGTATGGCAATAGAATNGTTCAAGA
ATTTTATGACCCCTATGACCCCAGTCAATGCCAGGTACGAATTGGTCAGGCTTNTTCACTG
GCTGGGCTGCTGCTTNTTCTGCCTTNTGGGAGGTGCCCTANTTGCTGTTCCCTGCGAAC

FIGURE 105

TCATAGGGGGCGCGATATTTCTTGCAGGTNTGGTTATTTAGTTGCCACAGCATGGTA
TGGCAATAGAATCGTTCAAGAATTNTATGACCTATGACCCAGTCAATGCCAGGTACGAAT
TTGGTCAGGCTCTNTTCACTGGNTGGGCTGCTGCTCTNTNNGCCTNTGGGAGGTGCCCTA
CTTGCTGTTCTG

FIGURE 106

TTCTGGATGGATCCGCCCCATCNTCACATGCCCTGCCCCNTGGAGATTACNCCTATGC
TGGCGAACACATCNTGACCGCCCAGGCCATGTACGAGGGCTGTGGAATGTCCTGCGTGTC
CCAGAGCACCAGGGCAGATCCAGTGCAAAGTCTTGACTCCTGCTGAATCTGAGCAGCACAT
TGCAAGCAACCNTGCCTTGATGGTGGTTGGCATCCTCCTGGAGTGATAGCAATCTTGTGG
CCACCGTTGGCATGAAAGTGTATGAAGTGCTTGGAAAGACGATGAGGTGCAGAAGATGAGGAT
GGCTGTCATTGGGGCGCGATATTCTTGCAGGTCTGGCTATTAGNNGCCACAGCAT
GGTATGGCAATCAGACCCNNTCANAAACTCTATGACCCATGACCCAGTCAATGCCAGGTA
CGAATTGGTCAGGCTCTTCACTGGCTGGCTGCTGCTCTGCCTCTGGGAGGTG
CCCTACTTTGCTGTTCTGTCCCCGAAAAACACCTTACCCACG

FIGURE 107

CGGGGCTGCAGCTGTTGGCTTCATCTCGCTTCTGGATGGAATCGGCCATCGTCAGCA
CTGCCCTGCCCATGGAGGATTTACTCNTATGCTGGCGACAACATCGTGACCNCAGGCCA
TGTACGAGGGCTGTTGATGTCNGCGTGTGCGAGAGCACCGGGCAGATCCAGTGCAAAGTCT
TTGACTCCTGCTGAATCTGAGCAGCACATTGCAAGCAACCNTGCCTGATGGTGGTTGGCA
TCCTCCTGGAGTGATAGCAATCTTGTGGCCACCGTTGGCATGAAGTGTATGAAGTGCTTG
GAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGCGCGATATTCTTCTTGC
AGGTCTGGCTATTNTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAAGAATTCTAT
GACCCTATGACCCCAGTCAATGCCAGGTACGAATTGGTCAGGCTCTCTTCACTGGCTGGC
TGCTGCTTCTCTGCCTCTGGGAGGTGCCCTACTTGTGTTCTGCGAA

FIGURE 108

GCCTGCCGTCAAGCTGCCGGGCACCGCGGCCCTGCCCTCGCCCTCGCCCTGCCCTGCAC
CGCGTAGACCGACCCCCCCCCTCCAGCGGCCACCCGGTAGAGGACCCCCGCCGTCCCCG
ACCGGTCCCCGCCTTTGTAAAACCTAAAGCGGGCGCAGCATTACGCTCCCGCCCCGGT
GACCTCTCAGGGTCTCCCCGCCAAAGGTGCTCGCCGCTAAGGAACATGGCGAAGGTGGAG
CAGGTCCCTGAGCCTCGAGCGCAGCACGAGCTCAAATTCCGAGGCTTCACCGATGTTGT
CACCAACCTAAAGCTGGCAACCCGACAGACGAAATGTGTGTTAAGGTGAAGACTA
CAGCACCACGTAGGTACTGTGTGAGGCCAACAGCGGAATCATCGATGCAGGGCCTCAATT
AATGTATCTGTGATGTTACAGCCTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTT
TATGGTTAGTCTATGTTGCTCCAAGTACACTTCAGATATGGAAGCAGTATGGAAGGAGG
CAAAACCGGAAGACCTTATGGATTCAAACCTTAGATGTGTGTTGAATTGCCAGCAGAGAAT
GATAAACACATGATGTAGAAATAAAATTATCCACAACACTGCATCAAAGACAGAAC
ACCAATAGTGTCTAAGTCTGAGTTCTTGGATGACACCAGAAGTTAAGAAGGTTATGG
AAGAATGTAAGAGGCTGCAAGGTGAAGTTAGAGGCTACGGGAGGAGAACAGCAGTTCAAG
GAAGAAGATGGACTGCGGATGAGGAAGACAGTCAGAGCAACAGCCCCATTTCAGCATTAGC
CCCAACTGGGAAGGAAGAAGGCCTAGCACCCGGCTTGGCTCTGGTGGTTGTTCTTA
TCGTTGGTGTAAATTGGGAAGATTGCCTTGAGAGGTAGCATGCACAGGATGGTAAATTG
GATTGGTGGATCCACCATATCATGGGATTAAATTATCATAACCATGTGTAAAAAGAAATT
AATGTATGATGACATCTCACAGGTCTGCTTAAATTACCCCTCCCTGCACACACATACAC
AGATACACACACAAATATAATGTAACGATCTTAGAAAGTTAAAATGTATAGTAACG
ATTGAGGGGGAAAAAGAATGATCTTATTAAATGACAAGGGAAACCATGAGTAATGCCACAAT
GGCATATTGTAAATGTCATTAAACATTGGTAGGCCTGGTACATGATGCTGGATTACCTC
TCTAAAATGACACCCTCCTCGCCTGGTGTGGCCCTGGGAGCTGGAGGCCAGCAT
GCTGGGGAGTGCCTGCAGCTCCACACAGTAGTCCCCACGTGGCCACTCCGGCCAGGCTG
CTTCCGTGCTTCAGTTCTGCAAGCCATCAGCTCCTGGACTGATGAACAGAGTCAGA
AGCCCAAAGGAATTGCACTGTGGCAGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGT
TGACTGATTGACCCAGCCTGGAAATAATGGCAGTGCTTGTACTTAAAGGGACCAA
GCTAAATTGTATTGGTCATGTAGTGAAGTCAAACGTGTTATTCAAGAGATGTTAATGCATA
TTAACTTATTAAATGTATTCATCTCATGTTCTTATTGTCACAAGAGTACAGTTAATGC
TGCCTGCTGCTGAACCTCTGTTGGTGAACGGTATTGCTGCTGGAGGGCTGTGGCTCCTCT
GTCTCTGGAGAGTCTGGTCATGTGGAGGTGGGTTATTGGGATGCTGGAGAACAGCTGCCA
GGAAGTGTGTTCTGGGTCACTAAATAACAACTGTCATAGGGAGGGAAATTCTCAGTAGTG
ACAGTCACACTAGGTTACCTTTAATGAAGAGTAGTCAGTCAGTCTCTAGATTGTTCTATA
CCACCTCTCAACCATTACTCACACTCCAGCGCCCAGGTCCAAGTCTGAGCCTGACCTCCCC
TTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAAGCGAGGGC
ACCAGCAGTGTGGTGGGGAGCAAGGGAAAGAGAGAACCTTCAGCGAATCCTCTAGTAC
TAGTTGAGAGTTGACTGTGAATTAAATTATGCCATAAAAGACCAACCCAGTTCTGTTGA
CTATGTAGCATCTTGAAGAAAAATTATAATAAGCCCCAAAATTAAAGAAAA

FIGURE 109

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977
<subunit 1 of 1, 243 aa, 1 stop
<MW: 27228, pI: 7.43, NX(S/T): 2
MAKVEQVLSLEPQHELKFRGPFTDVTTNLKLGNPDRNVCFKVKTAPRRYCVRPNSGIID
AGASINVSVMLQPFDYDPNEKSKHKFMVQSMFAPTDTSMDMEAVWKEAKPEDLMDSKLRCVFE
LPAENDKPHDVEINKIISTTAKTETPIVSKSLSSSLDDTEVKKMEECKRLQGEVQLRREE
NKQFKEEDGLRMRKTVQSNSPISALAPTGKEEGLSTRLLALVVLFFIVGVIIGKIAL
```

Important features:

Transmembrane domain:

amino acids 224-239

N-glycosylation site.

amino acids 68-71

N-myristoylation site.

amino acids 59-64, 64-69 and 235-240

FIGURE 110

GTCAGTCTTCTAGATTGTCCTTATCCCACCTTCAACCANTACTCACATTCNAGCGCCCAG
GTCCANGTCTGAGCCTGACTTCCCCTGGGGACCTAGCCTGGAGTCAGGACAATGGNTCGGG
CTGCAGAGGNTTAGAAGCGAGGGCACCAGCAGTTGGGTGGGGAGCAAGGGNNAGAGAAA
CTCTTCAGCGAATCCTTCTAGTACTAGTTGAGAGTTGACTGTGAATTAATTTATGCCATA
AAAGACNAACCCAGTTCTGTTGACTATGTAGCATCTGAAAAGAAAAATTATAATAAGCC
CCAAAATTAAGAATTCTTTGTCATTTGTCACATTCGCTATGGGGGAATTATTATTTT
ATCATTTTATTATTTGCCATTGGAAGGTTAACCTTAAAATGAGC

FIGURE 111

TATTGTAAAGGCCATTTAAACCATTGGTAGGCCTTGGTACATGATGCTGGATTACCTCCTT
AAATGACACCNTTCCTCGCCTGTTGGTCTGGCCNTGGGGAGCTGGAGCCCCAGCATGCTG
GGGAGTGCGGTCAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCGGCCCAAGGCTGCTT
CCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAACGCC
CAAAGGAATTGCCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGCGTGTGTTGA
CTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACTTAAAGGGACCAAGCT
AAATTGTATTGGTCATGTAGTGAAGTCAAACTGTTATTCAAGAGATGTTAATGCATATTTA
ACTTATTAAATGTATTCATCTCATGTTCTTATTGTCACAAGAGTACAGTTAATGCTGCG
TGCTGCTGAACCTGTTGGGTGAACGGTATTGCTGGAGGGCTG

FIGURE 112

CCCTGGTGGTTTGTCTTAATTGGTGTATTGGAAAGATTGCTTAGAGGTA
GNATGCACCNGGCTGGTAAATTGGATTGGATCCACCATATCCATGGGATTAAATTAT
CATAACCATGTGTAAAAGAAATTAATGTATGATGACATNTCACAGGTATTGCCTTAAATT
ACCCATCCCTGNANACACATACACAGATAACANANACAAATNTAATGTAACGATNTTTAG
AAAGTTAAAAATGTATAGTAAC

FIGURE 113

GGTGGCCCATTCCCGGCCAGGCTGCTTCCGGTNTTCAGTTCTGTCCAAGCCATCAGCTCC
TTGGGACTGATGAACAGAGTCAGAAGCCAAAGGAATTGCACTGTGGCAGCATNAGACGTAC
TTGTNATAAGTGAGAGGCGTGTGTTGACTGATTGACCCAGCGCTTGGAAATAAATGGCAGT
GCTTGTTCACTAAAGGGACCAAGCTAAATTGTATTGGTCATGTAGTGAAGTCAAACGT
TTATTCAGAGATGTTAATGCATATTAANTTAAATGTATTTNATNTCATGTTTCTTA
TTGTCACAAGAGTACAGTTAATGCTGCGTGCTGCTGAANTNTGTTGGGTGAACGTGGTATTGC
TGCTGGAGGGCTGTGGGCTCCTCTGTCTTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 114

TGCTTTCCGTGTCTTCAGTTCTGTCCAAGCCATCAGCTCCTGGGACTTGATGAACAGAGTC
AGAAGCCCAAAGGAATTGCACTGTGGCAGCATCAGACGTACTCGTCATAAGTGAGAGGGCGTG
TGTTGACTGATTGACCCAGCGCTTGGAAATAATGGCAGTGCTTGTTCACTTAAAGGGAC
CAAGCTAAATTGTATTGGTTCATGTAGTGAAGTCAAACTGTTATTAGAGATGTTAATGC
ATATTTAACCTATTTAATGTATTCATCTCATGTTCTTATTGTCACAAGAGTACAGTTAA
TGCTGCGTGC

FIGURE 115

AAACCTTAAAAGTTGAGGGAAAAGAATGATCCTTATTAATGACAAGGGAAACNTGN
AATGCCACAATGGCATATTGTAAATGTCATTAAACATTGGTAGGCCTGGTACATGATGC
TGGATTACCTCTCTTAAAATGACACCCTCCTCGCCTGTTGGTGTGGCCCTGGGGAGCTN
GAGCCCAGCATGCTGGGAGTGCAGCTGCTCCACACAGTAGTCCCCANGTGGCCANTCCC
GGCCCAGGCTGCTTCCGTGTCTCAGTTCTGTCCAAGCCATCAGCTCCTGGGANTGATGA
ACAGAGTCAGAAGCCAAAGGAATTGCANTGTGGCAGCATCAGANGTANTNGTCATAAGTGA
GAGGCGTGTGTGANTGATTGACCCAGCGCTTGGAAATAATGGCAGTGTGCTTGTTCANTT
AAAGGGNCCAAGNTAAATTGTATTGGTTCATGTAGTGAAGTCAAANTGTTATTCAAGAGATG
TTTAATGCATATTAAANTTATTAATGTATTCATNTCATGTTCTATTGTCACAAGGGT
ACAGTTAATGCTGCGTGTGCTGAANTCTGTTGGGTGAANTGGTATTGCTG

FIGURE 116

GGCCCTTGGGAGCTGGAGCCCAGCATGCTGGGAGTGCAGTCAGCTCCACACAGTAGTCCC
CACGTGGCCCCTCCGGCCAGGCTGCTTCCGTGTCTCAGTTCTGTCCAAGCCATCAGC
TCCTGGGACTGATGAACAGAGTCAGAAGCCAAAGGAATTGCACTGTGGCAGCATCAGACG
TACTCGTCATAAGTGAGAGCGTGTGTTGACTGATTGACCCAGCGCTTGAAATAATGGC
AGTGCTTGTTCACTTAAAGGGACCAAGCTAAATTGTATTGGTCATGTAGTGAAGTCAAA
CTGTTATTCAAGAGATGTTAATGCATATTTAACTTATTAAATGTATTCATCTCATGTTTC
TTATTGTCAAAAGAGTACAGTTAATGCTGCGTGTGCTGAACCTCTGTTGGGTGAACGGTAT
TGCTGCTGGAGGGCTGTGGCTCCTCTGTCTGGAGAGTCTGGTCATGTGGAGGTGGG

FIGURE 117

GGCAGCTCCGGTGCTGTGGCCGGCTTGGCGGGCGGCCTCCGGCTCAGGCTGGCTGAGA
GGCTCCCAGCTGCAGCGTCCCCGCCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCC
CTCGGGGACCAAACAAGCCTGGCAGGGTCTCACTTGTGCCCAGGCTGGAGTTAGTGCCTA
TGATCATGGTTACTGCAGCCTGACCTCCTGGGTTCAAGCGATCCTGCTGAGTAGCTGGG
CTACAGGACAAAATTAGAAGATCAAATGGAAAATATGCTGTTGGTTGATATTTTCACC
CCTGGGTGGACCCCTATTGATGGATCTGAAATGGAATGGATTATGTGGCACTTGAGAAA
GGTACCCGGATTGTCAGTGAAGGACTTCCATCTCACAGCCCCGATTGAGGCAGATG
CTAAGATGATGGTAAATACAGTGTGTCATCGAATGCCAGAAAGAACTCCAACTCCCAGC
CTTCTGAATTGGAGGATTATCTTCTATGAGACTGTCTTGAGAATGGCACCCGAACCTT
AACCAAGGGTGAAGATTCAAGATTGGTTCTTGAGCCGACTCAAATATCACCACAAAGGGAG
TATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCGACAGCAGGTTCAGCATCTTGAGAAA
AGGTTCTAACCAATTCCCTTCAGCACAGCTGTGAAGCTTCCACGGGTGAGTGGCAT
TCTCATTCCCTCAGCATGTTCTAACCTGCTGCCACTGTGTTCATGATGAAAGGACTATG
TCAAAGGGAGTAAAAGCTAACGGTAGGGTTGAGATGAGGAATAAAGTGGAGGCAAG
AAACGTCGAGGTTCTAACAGGGAGCAGGAGAGAAGCTAGTGGTGGTACCAAAGAGAGGGTAC
CAGAGAGCATCTGCAGGAGAGCAGAAGGGTGGAGAAGAAGAAAAAAATCTGGCCGGGTC
AGAGGATTGCCGAAGGGAGGCCCTTCAGGGACCCGGTCAAGAAATACCCACATTCCG
AAGGGCTGGGCACGAGGAGGCATGGGGACGCTACCTGGACTATGACTATGCTCTCTGGA
GCTGAAGCGTGTCAACAAAAGAAATACATGGAACCTTGGAAATCAGCCAAACGATCAAGAAA
TGCCTGGTGGAAATGATCCACTTCTCAGGATTGATAACGATAGGGCTGATCAGTTGGTCTAT
CGGTTTGCACTGTGTGTCGACGAATCCAATGATCTCCTTACCAATACTGCGATGCTGAGTC
GGGCTCCACCGTTGGGGTCTATCTCGTCTGAAAGATCCAGACAAAAGAATTGGAAGC
GCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGTTCAAGAAGGAC
TACAACGTTGCTGTCGATCACTCCCTAAAATACGCCAGATTGCTCTGGATTACCG
GAACGATGCCAACATTGTGCTAACGGCTAACAGAGACCTGAAACAGGGCGGTATCATCTAAA
TCACAGAGAAAACCAGCTCTGCTTACCGTAGTGAGATCACTCATAGGTATGCTGGACTT
GAACCTGTCAATAGCATTCAACATTTCAAAATCAGGAGATTTCGTCCTTAAAGACACACTTT
TGTATAGGTGCAGATATTGAAACTAGGTGGGACTTCAATGCCAAGTATAACTCTTCTTTA
CATGGTGATGAGTTCAATTGTAGAAAATTGTTGCCTCTTAAAGACACACTTT
AAACCTCAACAGGTATTATAAAACATGTGACTCCTTAATGGACTTATCTCAGGGTCC
TACTCTAAGAAGAATCTAATAGGATGCTGGTTGTATTAAATGTGAAATTGCAAGATAAA
GGTAGATGGTAAAGCAATTAGTATCAGAATAGAGACAGAAAGTTACAACACAGTTGTACTA
CTCTGAGATGGATCCATTCAAGCTCATGCCCTCAATGTTATATTGTGTTATCTGTTGGTCT
GGGACATTAGTTAGTTTTGAAGAATTACAAATCAGAAGAAAAGCAAGCATTATAAA
CAAAACTAATAACTGTTTACTGCTTAAGAAATAACAATTACAATGTGATTATTAAAAAA
TGGGAGAAAATAGTTGTTCTATGAAATAAACCTAGTTAGAAATAGGGAGCTGAGACATT
TAAGATCTCAAGTTTATTAACTAATACTCAAATATGGACTTTCATGTATGCAAGGG
AAGACACTTCACAAATTATGAATGATCATGTGTTGAAAGCCACATTATTTATGCTATACAT
TCTATGTATGAGGTGCTACATTAGGACAAAGAATTCTGTAATCTTCAAGAAAGAGT
CTTTCTCCTTGACAAAATCCAGCTTGTATGAGGACTATAGGGTGAATTCTCTGATTAG
TAATTAGATATGTCCTTCAAAATGAATAAAATTATGAATATGA

FIGURE 118

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVC
GIECQKELPTPSLSELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTGVSVRRKRQV
YGTDSRFSILDKRFLTNPFPSTAVKLSTGCSGILISPQHVLAAHCVHDGKDYVKGSKKLRV
GLLKMRNKSGGKKRGSKRSRREASGGDQREGTREHLQERAKGGRRKKSGRGQRIAEGRPS
FQWTRVKNTHIPKGWARGGMGDATLDYDYALLELKRAHKKYMELGISPTIKMPGMIHFS
GFDNDRADQLVYRFCVSDESNDLLYQYCDAESGSTGSGVYRLKDPKKNWKRKIIAVYSG
HQWVDVHGQKDYNNAVIRTPLKYAQICLWIHGNDANCAYG
```

Important features:

Signal peptide:

amino acids 1-16

N-glycosylation sites.

amino acids 90-93, 110-113 and 193-196

Glycosaminoglycan attachment site.

amino acids 236-239

Serine proteases, trypsin family, histidine active site.

amino acids 165-170

FIGURE 119

AATGTGAGAGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTAGCACCACTACTGGAT
GTGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCACTGTCCGATTCTGATTCCGGCAAGG
ATCCAAGCATGGAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTCTGGCTTTC
CTGCTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCTATGGGATGCCTG
GGGCCCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCTCTACTCTTGAGGCGCT
GCCTGAGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGAC
TGCCCCACCAGAAGCAGGTGATTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCA
CCATGGCCAGTTTATGAATGGCTCCTGTGTCTAATGACCCGTACAACCCATGTTCACTCA
AGTGCCAAGCAAAGGAACAAACCTGGTTGTTGAACCTAGCACCTAAGGTCTTAGATGGTACG
CGTTGCTATAACAGAAATCTTGGATATGTGCATCAGTGGTTATGCCAAATTGTTGGCTGCGA
TCACCAAGCTGGGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGTCCA
CCTGCCGGCTGGTCCGAGGGCAGTATAAATCCAGCTCCGCAACCAAATGGATGATACT
GTGGTTGCACCCATGGAAGTAGACATATTGCCCTGTCTTAAAGGTCTGATCACTT
ATATCTGGAAACCAAAACCCCTCCAGGGACTAAAGGTGAAACAGTCTCAGCTCCACAGGAA
CTTCCTTGTGGACAATTCTAGTGTGGACTTCCAGAAATTCCAGACAAAGAGATACTGAGA
ATGGCTGGACCACTCACAGCAGATTCATTGTCAAGATTGTAACCTGGCTCCGCTGACAG
TACAGTCCAGTCATCTTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTCTTTC
CTTGCTCAGCAACCTGTGGAGGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGG
AGCAACCGTGTGGTGCTGACCAAACTGTCAACTATTACCCAGAGAACATCAAACCAAACC
CAAGCTTCAGGAGTGCAACTGGATCCTGTCCAGCCAGTGACGGATAACAGCAGATCATGC
CTTATGACCTCTACCATCCCCCTCCTCGGTGGAGGCCACCCATGGACCGCGTGCTCCTCC
TCGTGTGGGGGGGCATCCAGAGCCGGCAGTTCTGTGTGGAGGGAGGACATCCAGGGCA
TGTCACTTCAGTGGAAAGAGTGGAAATGCATGTACACCCCTAACAGATGCCATCGCGAGCCCT
GCAACATTTTGACTGCCCTAAATGGCTGGCACAGGAGTGGTCTCCGTGACAGTGACATGT
GGCCAGGGCCTCAGATAACCGTGTGGCCTCTGCATCGACCATCGAGGAATGCACACAGGAGG
CTGTAGCCCCAAAACAAAGCCCCACATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATA
AACCCAAAGAGAAACTCCAGTCAGGGCCAAGTTGCCATGGTCAAACAAAGCTCAAGAGCTA
GAAGAAGGAGCTGCTGTGTCAAGAGGAGCCCTCGTAAGTTGTAAGACAGACTGTTCTATA
TTTGAAACTGTTTGTAAAGAAAGCAGTGTCTCACTGGTTGAGCTTCATGGGTTCTGA
ACTAAAGTGTAAATCATCTCACCAAGCTTTGGCTCTCAAATTAAAGATTGATTAGTTCAA
AAAAAAA

400128071042404

FIGURE 120

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847
<subunit 1 of 1, 525 aa, 1 stop
<MW: 58416, pI: 6.62, NX(S/T): 1
MECCRATPGTLLLFLAFLSSRTARSEEDRGLWDAGPWSECSRTCGGGASYSLRRCLS
SKSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQ
AKGTTLVVELAPKVLGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVNGDGSTCR
LVRGQYKSQLSATKSDDTVVALPYGSRHIRLVLKGPDHLYLETKTLQGTKGENSLSTGTFL
VDNSSVDFQKFPDKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFPCS
ATCGGGYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPASPASDGYKQIMPYD
LYHPLPRWEATPWTACSSSCGGGIQSRAVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNI
FDCPKWLAQEWSPTVTCGQGLRYRVVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPK
EKLPVEAKLPWFQQAQEEGAAVSEEPS
```

Important features:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 251-254

Thrombospondin 1

amino acids 385-399

von Willebrand factor type C domain proteins

amino acids 385-399, 445-459 and 42-56

FIGURE 121

CGGACGGCGTGGCGGCGCTCGGAACTCCCGTGGAGGGCCGTGGGCCCTCGGGCCTGAC
AGATGGCAGTGGCACTGCGGCGAGTACTGGCGCTCTGGCGGGCGCTGTGGCTGGCG
GCCCGCCGGTCTCGTGGGCCAGGGTCCAGCGGCTGCGCAGAGGCGGGACCCGGCCTCAT
GCACGGGAAGACTGTGCTGATCACCGGGCGAACAGCGGCCTGGCCGCCACGGCCCG
AGCTACTGCGCCTGGAGCGCGGGTGATCATGGCTGCCGGACCGCGCGCGCCAGGGAG
GCGGCGGGTCAGCTCCGCCGAGCTCCGCCAGGCCGGAGTGCAGGCCAGAGCCTGGCGT
CAGCGGGTGGCGAGCTCATAGTCCGGAGCTGGACCTCGCCTCGCTGCCTCGGTGCGCG
CCTTCTGCCAGGAAATGCTCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGG
ATCTTCCAGTGCCCTTACATGAAGACTGAAGATGGTTGAGATGCAGTCGGAGTGAACCA
TCTGGGCACTTCTACTCACCAATCTCCTTGGACTCCTCAAAAGTTCAAGCTCCCAGCA
GGATTGTGGTAGTTCTTCAAACCTTATAAAATACGGAGACATCAATTGATGACTTGAAC
AGTGAACAAAGCTATAATAAAAGCTTTGTTAGCCGGAGCAAAGTGGCTAACATTCTTT
TACCAAGGAACTAGCCGCCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTG
GTATTGTACGGACAAATCTGGGAGGCACATACACATTCCACTGTTGGTCAAACCACTCTTC
AATTGGTGTATGGCTTTTCAAACACTCCAGTAGAAGGTGCCAGACTCCATTATT
GGCCTCTCACCTGAGGTAGAAGGAGTGTCAAGGAAGATACTTGGGATTGTAAAGAGGAAG
AACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAAGTCTGGGATATCAGTGAAGTG
ATGGTTGGCCTGCTAAAATAGGAACAAGGAGTAAAGAGCTGTTATAAAACTGCATATCAG
TTATATCTGTGATCAGGAATGGTGTGGATTGAGAACTGTTACTTGAAGAAAAGAATTGG
ATATTGGAATAGCCTGCTAACAGGTACATGTGGGTATTTGGAGTTACTGAAAAATTATTT
TGGGATAAGAGAATTTCAGCAAAGATGTTAAATATATAGTAAGTATAATGAATAATAA
GTACAATGAAAAATACAATTATATTGAAATATAACTGGCAAGCATGGATGACATATTA
ATATTGTCAGAATTAAAGTGAATCAAAGTGCTATCGAGAGGTTTCAAGTATCTTGAGTT
TCATGGCCAAAGTGTAAACTAGTTTACTACAATGTTGGTGTGGAAATTATCTGC
CTGGTGTGTGCACACAAGTCTTACTTGAATAAATTACTGGTAC

FIGURE 122

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747
<subunit 1 of 1, 336 aa, 1 stop
<MW: 36865, pI: 9.15, NX(S/T): 2
MAVATAAAVLAALGGALWLAARRFVGPRVQQLRRGGDPGLMHGKTVLITGANSGLGRATAAE
LLRLGARVIMGCRDRARAEEAAGQLRRELRQAAECGPEPGVSGVGELIVRELDLASLRSVRA
FCQEMLQEEPRLDVLINNAGIFQCPYMKTEDGFEMQFGVNHLGHFLLTNLLGLLKSSAPSR
IVVVSSKLYKYGDINFDDLNEQSYNKSFCYSRSKLANILFTRELARRLEGTNVTVNVLHPG
IVRTNLGRHIHPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSGRYFGDCKEEE
LLPKAMDESVARKLWDISEVMVGLLK
```

Important features:

Signal peptide:

amino acids 1-21

Short-chain alcohol dehydrogenase family protein

amino acids 134-144, 44-56 and 239-248

N-glycosylation site.

amino acids 212-215 and 239-242

FIGURE 123

GGGGATTGTAAAGAGGAAGNACTGTGCCCAAAGNTATGGATGAATCTGTTGCAAGAAAATTN
TGGGATATCAGTGAAGTGTGGTTNGCCTGCTAAAATAGGAACAAGGAGTAAAAGAGCTGTT
TATAAAACTGCATATCAGTTATCTGTGATCAGGAATGGTGTGGATTGAGAACTTGTACT
TGAAGAAAAAGAATTTGATATTGGAATAGCCTGNTAAGAGGNACATGTGGTATTGGAG
TTACTGAAAAATTATTTGGGATAAGAGAATTCAGCAAAGATGTTAAATATATAGT
AAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTTAAATTATAACTGGCA
AGCATGGATGACATATTAATATTGTCAGAATTAAAGTGAATCAAAGTGCTATCGAGAGGTT
TTCAAGTATCTTGAGTTCATGCCAAAGTGTAACTAGTTTACTACAATGTTGGTGT
TGTGTGGAAATTATCTGCCTGGCTT

FIGURE 124

GAGAGGACGAGGTGCCGCTGCCCTGGAGAACCTCCGCTGCCGTCGGCTCCGGAGGCCAGCC
CTTCCTAACCAACCAACCTAGCCCAGTCCCAGCCAGCGCCTGTCCTGTCACGGAC
CCCAGCGTTACCATGCATCCTGCCGTCTCCTATCCTAACCGACCTCAGATGCTCCCTCT
GCTCCTGGTAACGGGTTTTACTCCTGTAACAACGTGAAATAACAAGTCTTGCTACAGAGA
ATATAGATGAAATTAAACAATGCTGATGTTGCTTAGTAAATTATGCTGACTGGTGT
CGTTTCAGTCAGATGTTGCATCCAATTTGAGGAAGCTCCGATGTCATTAAGGAAGAATT
TCCAAATGAAAATCAAGTAGTGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCC
AGAGATAACAGGATAAGCAAATACCAACCCCTCAAATTGTTGTAATGGGATGATGATGAAG
AGAGAATAACAGGGGTCAAGCGATCAGTGAAAGCATTGGCAGATTACATCAGGCAACAAAAAG
TGACCCCCATTCAAGAAATTGGACTTAGCAGAAATCACCCTTGATCGCAGCAAAAGAA
ATATCATTGGATATTTGAGCAAAAGGACTCGGACAACATAGAGTTTGAACGAGTAGCG
AATATTTGCATGACTGTGCCTTCTGCATTGGGATTTCAAAACCGGAAAG
ATATACTGGCGACAACATAATCTACAAACCACCAGGGCATTCTGCTCCGGATATGGTGTACT
TGGGAGCTATGACAAATTGATGTGACTTACAATTGGATTCAAGATAATGTGTTCCCTCTT
GTCCGAGAAATAACATTGAAAATGGAGAGGAATTGACAGAAGAAGGACTGCCCTTCTCAT
ACTCTTCACATGAAAGAAGATAACAGAAAGTTAGAAATATTCCAGAAATGAAGTAGCTCGGC
AATTAAATAAGTAAAAAGGTACAATAAACCTTTACATGCCGATTGTGACAAATTAGACAT
CCTCTCTGCACATACAGAAAACCTCCAGCAGATTGTCCTGTAATCGCTATTGACAGCTTAG
GCATATGTATGTGTTGGAGACTTCAAAGATGTATTAATTCCGGAAAACCAAGCAATTG
TATTTGACTTACATTCTGGAAAACGCACAGAGAATTCCATCATGGACCTGACCCAACTGAT
ACAGCCCCAGGAGAGCAAGCCAAAGATGTAGCAAGCAGTCCACCTGAGAGCTCCTCCAGAA
ACTAGCACCCAGTGAATATAGGTATACTCTATTGAGGGATCGAGATGAGCTTAAAAACTTG
AAAAACAGTTGTAAGCCTTCAACAGCAGCATCAACCTACGTGGTGGAAATAGTAAACCTA
TATTTCTATAATTCTATGTGTATTTTATTTGAATAACAGAAAGAAATTAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 125

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57689
<subunit 1 of 1, 406 aa, 1 stop
<MW: 46927, pI: 5.21, NX(S/T): 0
MHPAVFLSLPDLRCSLLLLTVWVFTPVTTEITSLATENIDEILNNADVALVNFYADWCRFSQ
MLHPIFEEASDVIKEEFPNENQVVFARVDCDQHSDIAQRYRISKYPTLKLFRNGMMMKREYR
GQRSVKALADYIRQQKSDPIQEIRDLAETTLDRSKRNIIGYFEQKDSNDNYRVFERVANILH
DDCAFLSAFGDVSKEPERYSGDNIIYKPPGHSAPDMVYLGAMTNFDVTYNWIQDKCVPLVREI
TFENGEELTEEGLPFLILFHMKEDTESLEIFQNEVARQLISEKGTINFLHADCDKFRHPLLH
IQKTPADCPVIAIDSFRHMYVFGDFKDVLIPGKLQFVFDLHSGKLHREFHHGPDPTDTAPG
EQAQDVASSPPESSFQKLAPSEYRYTLLRDRDEL
```

Important features:

Signal peptide:

amino acids 1-29

Endoplasmic reticulum targeting sequence.

amino acids 403-406

Tyrosine kinase phosphorylation site.

amino acids 203-211

Thioredoxin family proteins

amino acids 50-66

FIGURE 126

ATTAAGGAAGAATTCCAAATGAAAATCAAGTAGTNTTGCCAGAGTNGATTGTGATCAGCA
CTCTGACATAGCCCAGAGATAACAGGATAAGCAAATACCCAAACCTCAAATTGTTCGTAATG
GGATGATGATGAAGAGAGAATACAGGGGTCAAGCGATCAGTGAAGCATGGCAGATTA

FIGURE 127

AGAGGCCTCTGGAAAGTTGCCGGGTGTCGCCGNGAGCCGGTCGAGAGGACNAGG
TGCCGCTGCCGGAGAACCTCCGCTGCCGTCGGCTCCGGAGCCCAGCCCTTCCTAACCC
AACCCAACCTAGCCCNGTCCCAGCCGCCAGCGCCTGTCCTGTCNCGGANCCAGCGTNACC
ATGCATCCTGCCGTCTCCTATCCTTACCCGACCTCAGATGCTCCCTGCTCCTGGTAAC
TTGGGTTTTACTCCTGAAACAACGTAAACNNGTCTTGATACNNAGAATATAGATGAAA
TTTAAACNATGCTGATGTGGCTTAGTCATTTATGCTGACTGGTGTCTTCAGTCAG
ATGTGGCATCCAATTTTGAGGANGCTCCGATGTCATTAAGGAAGAATTCCAATGAAAA
TCAAGTAGTGTTGCCAGAGTTGATTGTGATCAGCACTCTGACATAGCCCAGAGATAACAGGA
TAAGCAAATACCCAACCCTCAAATTGTTGTAATGGGATGATGATGAAGAGAGAATAACAGG
GGTCAGCGATCAGTGAAGCATTGGCAGATTACATCAGGC

卷之三

FIGURE 128

CCCCACGCGTCCG**AT**GGCGTTACGTTCGGGCCTCTGCTACATGCTGGCGCTGCTGCTCA
CTGCCGCGCTCATCTTCTCGCATTGGCACATTATAGCATTGATGAGCTGAAGACTGAT
TACAAGAATCCTATAGACCA**G**GTAA**T**ACCC**T**GAATCCC**T**TGACTCCCAGAGTAC**T**CAT
CCACGCTTCTTGT**C**ATGTTCTTGTGCAGCAGAGTGGCTTACACTGGGTCTCAATA
TGCCCC**T**GGCATAT**C**ATATTGGAGGTATATGAGTAGACCA**G**GTGAGTGGCC**C**AGGA
CTCTATGAC**C**CTACAA**C**CCATCATGAATGCAGATATTCTAGCATATTGTCAGAAGGAAGGATG
GTGCAAATTAGCTTTATCTTCTAGCATTTTTACTACCTATATGGCATGATCTATGTT
TGGTGAGCT**C****T****A**GA**A**ACACACAGAAGAATTGGTCCAG**T**A**G**TGCATGC**AAA**AGCCAC
CAAATGAAGGGATTCTATCCAGCAAGATCCTGTCC**A**AG**G**TAGCCTGTGG**A**CTGATCAGT
TACTTT**AAA**ATGACTCCTTATTTTAAATGTTCCACATTTGCTTG**G**AAAGACTG
TTTCATATGTTACTCAGATAAAGATTAA**G**GTATTACGTATA**A**TTAATATA**AA**AT
GATTACCTCTGGTGTGACAGGTTGA**A**CTTGCACTTCTTA**A**GGAA**C**AGCCATAATCCTCTG
AATGATGCATTAATTACTGACTGTCCTAGTACATTGG**A**GCTTTGTTATAGGA**A**CTTGTA
GGGCTCATTGGTTCAT**G**AA**C**AGTAT**C**TAATTATA**AA**ATTAGCTGTAGATATCAGGTGC
TTCTGATGA**A**GT**G**AA**A**GTATATCTGACTAGTGG**G**AA**A**CTTCATGGTTCC**C**TCATCTGTC
ATGTCGATGATTATATGGATA**C**ATT**A**AAA**A**AGCGGG**A**TTTCC**C**TCATCGCTT
GAATATTATCC**C**GTATATTGCAT**G**AA**T**GAGAGATT**T**CCC**A**ATT**C**CAT**G**AGTA**A**AAA
TATACTTGCTTAA**T**CTTA**A**AGC**A**TA**G**TA**A**ACATG**A**T**A**AAA**A**ATATGCTGAATTACTT
GTGA**A**GA**A**ATGC**A**TT**A**AG**C**TATTTAA**G**TGTTTATTGTA**A**GACATTACTTATT**A**AG
AAATTGGTTATTATGCTTACTGTT**C**TAATCTGGTGG**T**AA**AG**GTATTCTTA**A**GAATT**T**GCAGG
TACTACAGATTTC**AAA**ACTG**A**ATGAGAG**A**AAATT**G**TATAACC**C**C**T**GCTGTT**C**TTAGT
GCA**A**T**C**A**A**AAA**A**CTG**C**AA**T**TA**A**AG**A**CTC

FIGURE 129

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23330
<subunit 1 of 1, 144 aa, 1 stop
<MW: 16699, pI: 5.60, NX(S/T): 0
MAFTFAAFCYMLALLTAALIFFAIWHIIIAFDELKTDYKNPIDQCNLNPVLPEYLIHAFF
CVMFLCAAELTLGLNMPLLAYHIWRYMSRPVMSGPGLYDPTTIMNADILAYCQKEGWCKLA
FYLLAFFYYLYGMIYVLVSS
```

Important features:

Signal peptide:

amino acids 1-20

Type II transmembrane domain:

amino acids 11-31

Other transmembrane domain:

amino acids 57-77 and 123-143

FIGURE 130

ATTATAGCATTGATGAGCTGAAGACTGATTACAAGATCCTATAGACCAGTGTAAATACCTG
AATCCCCTTGTACTCCCAGAGTACCTCATCCACGCTTCTTGTGTATGTTCTTGTGC
AGCAGAGTGGCTTACACTGGGTCTCAATATGCCCTCTGGCATATCATATTTGGAGGTATA
TGAGTAGACCAGTGATGAGTGGCCCAGGACTCTATGACCTACAACCATCATGAATGCAGAT
ATTCTAGCATATTGTCAGAAGGAAGGATGGTCAAATTAGCTTTATCTTCTAGCATT
TTACTACCTATATGGCATGATCTATGTTGGTGAGCTCTAGAACACACAGAAGAATT
GGTCCAGTTAAGTGCATGCAAAAGCCACAAATGAAGGGATTCTATCCAGCAAGATCCTGT
CCAAGAGTAGCCTGTGGAATCTGATCAGTTACTTTAAAAAATG

FIGURE 131

CGGACCGCGTGGGGAAACCCCTCCGAGAAAACAGCAACAAGCTGAGCTGCTGTGACAGAGGG
GAACAAGATGGCGCGCCGAAGGGAGCCTCTGGGTGAGGACCCAACTGGGCTCCGCCGC
TGCTGCTGCTGACCATGCCCTGCCGGAGGTCGGGGACCGCTCGCTGAAGCATTGAC
TCGGTCTGGGTGATACGCGTCTGCCACCGGGCTGTCAGTTGACCTACCCCTGCACAC
CTACCCTAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTC
AGTTTGTGGATGATGGAATTGACTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACA
GAAGCATATTCCAATCTGATGAGCAATATGCTGCCATCTGGTGCAGAACAGCTGCC
ATTCGCTGAACTGAGACAACAACTTATGTCCTGATGCCAAAATGCACCTACTCTTC
CTCTAACTCTGGTGAGGTCATTCTGGAGTGACATGATGGACTCCGCACAGAGCTTCATAACC
TCTTCATGGACTTTTATCTTCAAGCCATGACGGAAAAATAGTTATATTCCAGTCTAAGCC
AGAAAATCCAGTACGCACCACATTGGAGCAGGAGCCTACAAATTGAGAGAACATCTCTAA
GCAAAATGTCCTATCTGCAAATGAGAAATTCACAGCGCACAGGAATTCTTGAAGATGGA
GAAAGTGATGGCTTTAAGATGCCTCTCTTAACTCTGGGTGGATTAACTACAACTCT
TGTCCCTCGGTGATGGTATTGCTTGGATTGTTGCAACTGTTGCTACAGCTGTGGAGC
AGTATGTTCCCTCTGAGAAGCTGAGTATCTATGGTACTTGGAGTTATGAATGAACAAAAG
CTAAACAGATATCCAGCTTCTCTTGTGGTTGTTAGATCTAAACTGAAGATCATGAAGA
AGCAGGGCCTTACCTACAAAGTGAATCTGCTCATTCTGAAATTAAAGCATTTCTTT
AAAAGACAAGTGAATAGACATCTAAATTCCACTCCTCATAGAGCTTTAAAATGGTTCA
TTGGATATAGGCCTTAAGAAACTAAAGTTACTCAAATCTGTG

40012023-1207-1404

FIGURE 132

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA26847
<subunit 1 of 1, 323 aa, 1 stop
<MW: 36223, pI: 5.06, NX(S/T): 1
MAAPKGSLWVRTQLGLPPLLLTMALAGGSGTASAEAFDSVLGDTASCHRACQLTYPLHTYP
KEEELYACQRGCRLFSICQFVDDGIDLNRTKLECESACTEAYSQSDEQYACHLGCQNQLPFA
ELRQEQLMSLMPKMHLLFPLTLVRSFWSDMMDSAQSFITSSWTFYLQADDGKIVIFQSKPEI
QYAPHLEQEPTNLRESSLSKMSYLQMRNSQAHRNFLEDGESDGFLRCLSLNSGWILTTLVL
SVMVLLWICCATVATAVEQYVPSEKLSIYGDLEFMNEQKLNRYPASSLVVRSKTEDHEEAG
PLPTKVNLAHSEI
```

Important features:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 241-260

N-glycosylation site.

amino acids 90-93

FIGURE 133

TTGGGTGATA CGCGTCTTGCACCGGGCCTGTCAGTTGACCTACCCCTGCACACCTACCC
TAAGGAAGAGGAGTTGTACGCATGTCAGAGAGGTTGCAGGCTGTTCAATTGTCAGTTG
TGGATGATGGAATTGACTTAAATCGAACTAAATTGGAATGTGAATCTGCATGTACAGAAGCA
TATTCCAATCTGATGAGCAATATGCTGCCATCTGGTTGCCAGAACATCAGCTGCCATTGC
TGAACTGAGACAAGAACAACTTATGTCCCTGATGCCAAAATGCACCTACTCTTCCTCTAA
CTCTGGTGAGGTCAATTCTGGAGTGACATGATGGACTCCGC

FIGURE 134

CACACTGGCCGGATCTTTAGAGTCCTTGACCTTGACCAAGGGTCNGGAAAACAGCAACAA
GCTGAGCTGCTGTGACAGAGGGAAACAAGATGGCGGCCGAAGGGAGCCTTGGGTGAGGAC
CCAATGGGGCTCCCGCCGCTGCTGCTGACCATGGCCTGGCCGGAGGTTCGGGGACCG
CTTCGGCTGAAGCATTGACTCGGTCTTGGGTGATAACGGCGTCTGCCACCGGGCTGTCAG
TTGACCTACCCCTTGACACCTACCCCTAACCGAAGAGGGAGTTGTACGCATGTCAGAGAGGTTG
CAGGCTGTTTCAATTGTCAGTTGTGGATGATGGAATTGACTTAAATCGAACTAAATTGG
AATGTGAATCTGCATGTACAGAACATATTCCAATCTGATGAGCAATATGCTGCCATCTT
GGTTGCCAGAACATCAGCTGCCATTGCTGAACGTGAGACAAGAACAACTTATGTCCCTGATGCC
AAAAATGCACCTACTCTTCCTCTAACTCTGGTGAGGTCTTCTGGAGTGACATGATGGACT
CCGC

FIGURE 135

GCGAGGTGGCGATCGCTGAGAGGCAGGAGGGCGAGGCAGGGCGGGCCTGGGAGGCAGGCCGGAGGT
GGGGCGCCGCTGGGCCGGCCGCACGGGCTTCATCTGAGGGCGCACGGCCCGCACCGAGC
GTGCGGACTGGCCTCCCAAGCGTGGGCCGACAAGCTGCCGGAGCTGCAATGGGCCGGCTG
GGGATTCTTGTGTTGGCCTCCTGGCGCCGTGTGGCTGCTCAGCTCGGCCACGGAGAGGAGC
AGCCCCGGAGACAGCGGCACAGAGGTGCTCTGCCAGGTTAGTGGTTACTTGGATGATTGT
ACCTGTGATGTTGAAACCATTGATAGATTAAATAACTACAGGCTTTCCAAGACTACAAAA
ACTTCTGAAAGTGA~~CT~~ACTTTAGGTATTACAAGGTAACCTGAAGAGGCCGTGTCCTTCT
GGAATGACATCAGCCAGTGTGGAAGAAGGGACTGTGCTGTCAAACCATGTCAATCTGATGAA
GTTCCTGATGGAATTAAATCTGCGAGCTACAAGTATTCTGAAGAAGCCAATAATCTCATTGA
AGAATGTGAACAAGCTGAACGACTTGGAGCAGTGGATGAATCTCTGAGTGAGGAAACACAGA
AGGCTGTTCTCAGTGGACCAAGCATGATGATTCTCAGATAACTCTGTGAAGCTGATGAC
ATTCA~~G~~GTCCCCTGAAGCTGAATATGTAGATTGCTTCTTAATCCTGAGCGCTACACTGGTTA
CAAGGGACCAGATGCTGGAAAATATGGAATGTCATCTACGAAGAAA~~ACT~~GTTAAGCCAC
AGACAATTAAAAGACCTTAAATCCTTGGCTCTGGTCAAGGGACAAGTGAAGAGAACACT
TTTACAGTTGGCTAGAAGGTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGG
CCTACATGCAAGCATTAAATGTGCATTGAGTGCAAGATATCTTACAAGAGACCTGGTTAG
AAAAGAAATGGGGACACAAACATTACAGAATTCAACAGCGATTGATGGAATTGACTGAA
GGAGAAGGTCCAAGAAGGCTTAAGAACTTGTATTTCTACTTAATAGAACTAAGGGCTTT
ATCCAAAGTGTACCATTCTCGAGCGCCAGATTCAACTCTTACTGGAAATAAAATTC
AGGATGAGGAAAACAAATGTTACTTCTGAAATACTTCATGAAATCAAGTCATTCTTGTG
CATTTGATGAGAATTCA~~TTT~~GCTGGGATAAAAAGAAGCACACAAACTAAAGGAGGA
CTTTCGACTGCATTAGAAATATTCAGAAATTATGGATTGTGTTGTTAAATGTC
GTCTGTGGGAAAGCTCAGACTCAGGGTTGGCACTGCTCTGAAGATCTTATTTCTGAG
AAATTGATAGCAAATATGCCAGAAAGTGGACCTAGTTATGAATTCCATCTAACCAAGACAAGA
AATAGTATCATTATTCAACGCATTGGAAGAATTCTACAAGTGTGAAAGAATTAGAAA~~ACT~~
TCAGGAAC~~T~~GTACAGAATATTCATTAAAGAAAACAAGCTGATATGTGCCTGTTCTGGAC
AATGGAGGCAGAGGTGGAATTTCATTCAAAGGCATAATAGCAATGACAGTC~~TTAAGCCAA~~
ACATT~~TT~~TATAAAAGTTGCTTTGTAAAGGAGAATTATATTGTTAAGTAAACACATT~~TT~~
AAAATTGTGTTAAGTCTATGTATAACTACTGTGAGTAAAAGTAATACTTAAATAATGTG
GTACAAATTAAAGTTAATATTGAATAAAAGGAGGATTATCAAATTAAAAAAA
AAAAAAAAAAAAAAA

10012032402404

FIGURE 136

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53974
<subunit 1 of 1, 468 aa, 1 stop
<MW: 54393, pI: 5.63, NX(S/T): 2
MGRGWGFLFGLLGAVWLSSGHGEEQPETAAQRCFCQVSGYLDDCTCDVETIDRFNNYRLF
PRLQKLLESDYFRRYYKVNLRPCPFWNDISQCGRRDCAVKPCQSDEVPDGKSYKSEEAA
NNLIEECEQAERLGAVDESLSEETQKAVLQWTKHDDSSDNFCEADDIQSPEAEYVDLLNPE
RYTGYKGPDAWKIWNVIYEENCFKPQTIKRPLNPLASGQGTSEENTFYSWLEGLCVEKRAFY
RLISGLHASINVHLSARYLLQETWLEKKWGHNITEFQQRFDGILTEGEGPRLKNLYFLYLI
ELRALSKVLPFFERPDFQLFTGNKIQDEENKMLLEILHEIKSFPLHFDENSFFAGDKKEAH
KLKEDFRLHFRNISRIMDCVGCFKCRLWGKLOQGLGTALKILFSEKLIANMPESGPSYEFH
LTRQEIVSLFNAFGRISTSVKELENFRNLLQNIH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 280-283 and 384-387

Amidation site.

amino acids 94-97

Glycosaminoglycan attachment site.

amino acids 20-23 and 223-226

Aminotransferases class-V pyridoxal-phosphate

amino acids 216-222

Interleukin-7 proteins

amino acids 338-343

FIGURE 137

GCTGGAAATATGGATGTCATCTACGAGAACTGTTTAAGCCACAGACAATTAAAAGACCTT
TAAATCCTTGCTTCTGGTCAAGGGACAAGTGAAGAGNACACTTTTACAGTTGGCTAGAA
GGTCTCTGTGTAGAAAAAGAGCATTCTACAGACTTATATCTGGCCTACATGCAAGCATTAA
TGTGCATTTGAGTGCAAGATATCTTTACAAGAGACCTGGTTAGAAAAGAAATGGGGACACA
ACATTACAGAATTNAACAGCGATTGATGGAATTGACTGAAGGAGAAGGTCCAAGAAGG
CTTAAGAACTTGTATTTCTCTACTTAATAGAACTAAGGGCTTATCCAAAGTGTACCTT
CTTNGAGCGCCCAGATTTCAACTNTTACTGGAAATAAAATTCAAGGATGAGGNAAACAAAA
TGTTACTTTGGAAATACTTCATGAAATCAAGTCATTCCTTGCAATTGATGAGAATTCA
TTTTTTGCTG

FIGURE 138

CGGACGGCGTGGCGGACCGCGTGGCGGACCGCGTGGGTGGGAGGGGGCAGGATGGGAGGGAA
AGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAC
CGATCAGGCATGGAACTCCCCTTCGTCACTCACCTGTTCTGCCCTGGTGTCCCTGACAGG
TCTCTGCTCCCCTTAACCTGGATGAACATCACCCACGCCTATTCCAGGGCCACCAGAAG
CTGAATTGGATACAGTGTCTTACAACATGTTGGGGTGGACAGCGATGGATGCTGGTGGC
GCCCTGGATGGCCTTCAGCGACCGGAGGGGGACGTTATCGCTGCCCTGTAGGGGG
GGCCACAATGCCCATGTGCCAAGGCCACTTAGGTGACTACCAACTGGAAATTCATCTC
ATCCTGCTGTGAATATGCACCTGGGATGTCTCTGTTAGAGACAGATGGTATGGGGATT
ATGGTGAGCTAAGGAGAGGGTGGCAGTGTCTGAAGGTCCATAAAAGAAAAAGAGAA
GTGTGTAAGGAAAATGGTCTGTGTGGAGGGTCAAGGAGTTAAAAACCTAGAAAGCAA
AGGTAGGTAATGTCAGGGAGTAGTCTTCATGCCTCCTCACTGGGAGCATGTTCTGAGGGT
GCCCTCCAAGCCTGGAGTAACTATTCCCCCATCCCCAGGCCTGTGCCCTCTGGTCT
CGTGCTTGTGGCAGCTGTCTTCAGTTCTGGATATGTGCCGTGTGGATGCTTCATTCCA
GCCTCAGGAAGCCTGGCACCCACTGCCAACGTGAGCCAGAGGAAGGCTGAGTACTGGTT
CCCAGAAGGAGATACTGGTGGAAAAAGATGGGCAAAGCGGTATGATGCCCTGGCAAAGGG
CCTGCATGGCTATCCTCATTGCTACCTAACGCTGCTGCAAAGCTCCATGTTCTAACAGA
TTCAGACTCCTGGCCAGGTGTGGCCACACCTGTAATTCTAGCAGCTTGGGAGGCCAAG
GTGGCAGATCACTTGAGGTCAAGGAGTCAAGACCAGCCTGGCAAACATGGTAAACTCCAT
CTCTACTAAAAAAAAAAACAAAAATTAGCTGGTGCCTAGTCATGCCGTAAATCTC
ATCTACTCGGGAGGCTAACAGACAGGAGACTCTCACTAACCCAGGAGGTGGAGGTTGCCGTG
AGCCAAGATTGTGCCCTGCACTCTAGCGTGGGTGACAGAGTAAGCGAGACTCCATCTCAA
AATAATAATAATAATTCAACTCCTTATCAGGAGTCCATGATCTGGCCTGGCACAGTAA
CTCATGCCGTAAATCCAACATTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGAG
GTTTGAGACCAGCCTGGCAACATAGAAAGACCCATCTCTAAATAATGTTAAAAAT

FIGURE 139

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSPFNLDDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPW
DGPSGDRRGDVYRCPVGGAHNAPCAKGHLGDYQLGNSSHAPAVNMHLGMSLLETGDGGFMVS
```

Important features:

Signal peptide:

amino acids 1-22

Cell attachment sequence.

amino acids 70-73

N-glycosylation site.

amino acids 98-101

Integrins alpha chain proteins

amino acids 67-81

FIGURE 140

CACAGTTCCCCACCATCACTCNCCCATTCTTCCAACTTATTTAGCTTGCCTGGGA
GGGGGCAGGATGGGAGGGAAAGTGAAGAAAACAGAAAAGGAGAGGGACAGAGGCCAGAGGAC
TTCTCATACTGGACAGAAACCGATCAGGCATGGAACCTCCCTCGTCACTCACCTGTTCTG
CCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTAACCTGGATGAACATCACCCACGCCT
ATTCCCAGGGCCACCAGAAGCTGAATTGGATAACAGTGTCTAACACATGTTGGGGTGGAC
AGCGATGGATGCTGGTGGCGCCCCCTGGATGGCCTTCAGGCACCGGAGGGGGACGTT
TATCGCTGCCCTGTAGGGGGGGCCCACAATGCCCATGTGCCAACGGCCACTTAGGTGACTA
CCAACCTGGAAATTCATCTCATCCTGCTGTGAATATGCACCTGGGATGTCTCTGTTAGAGA
CAGATGGTGTGG

FIGURE 141

AAAGTTACATTTCTCTGGAACCTCTCCTAGGCCACTCCCTGCTGATGCAACATCTGGGTTTG
GGCAGAAAGGAGGGTGCCTCGGAGCCGCCCTTCTGAGCTCCTGGGCCGGCTCTAGAACAA
ATTCAAGGCTTCGCTGCGACTCAGACCTCAGCTCCAACATATGCATTCTGAAGAAAGATGGCT
GAGATGGACAGAATGCTTATTTGGAAAGAACAAATGTTCTAGGTCAAACGTGAGTCTACCA
AATGCAGACTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCTTTCATGTGGTTTTCT
ACGCATTGATTCCATGTTGCTCACAGATGAAGTGGCCATTCTGCCTGCCCTCAGAACCTC
TCTGTACTCTCAACCAACATGAAGCATCTTGATGTGGAGCCCAGTGATCGCGCTGGAGA
AACAGTGTACTATTCTGTCGAATACCAGGGGAGTACGAGAGCCTGTACAGAGCCACATCT
GGATCCCCAGCAGCTGGTCTCACTCACTGAAGGCTTGAGTGTGATGTCAGTGACATC
ACGGCCACTGTGCCATACAACCTTCGTGTCAGGGCACATTGGGCTCACAGACCTCAGCCTG
GAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCACCTTACCCGACCTGGGATGGAGA
TCACCAAAGATGGCTTCCACCTGGTTATTGAGCTGGAGGACCTGGGGCCAGTTGAGTTC
CTTGTGGCCTACTGGAGGAGGGAGCCTGGTGCAGGAAACATGTCAAAATGGTGAGGAGTGG
GGGTATTCCAGTGCACCTAGAAACCATGGAGCCAGGGCTGCATACTGTGTGAAGGCCAGA
CATTGTAAGGCCATTGGGAGGTACAGCGCCTTCAGCCAGACAGAAATGTGTGGAGGTGCAA
GGAGAGGCCATTCCCCTGGTACTGGCCCTGTTGCCTTGTGCTCATGCTGATCCTTGT
GGTCGTGCCACTGTTGCTGGAAAATGGGCCGGCTGCTCCAGTACTCCTGTTGCCCGTGG
TGGTCCTCCCAGACACCTTGAAAATAACCAATTCACCCAGAAGTTAATCAGCTGCAGAAGG
GAGGAGGTGGATGCCTGTGCCACGGCTGTGATGTCCTGAGGAACCTCTCAGGGCTGGAT
CTCA**TAG**TTGCGGAAGGGCCAGGTGAAGCCAGAACCTGGTCTGATGACATGGAAACC
ATGAGGGACAAGTTGTTCTGTTCCGCCACGGACAAGGGATGAGAGAAGTAGGAAGA
GCCTGTTGCTACAAGTCTAGAACCAACCATCAGAGGCAGGGTGGTTGTCTAACAGAACAC
TGACTGAGGCTTAGGGGATGTGACCTCTAGACTGGGGCTGCCACTGCTGGCTGAGCAACC
CTGGGAAAAGTGACTTCATCCCTCGGTCTAACGTTCTCATCTGTAATGGGGAAATTACC
TACACACCTGCTAACACACACACAGAGTCTCTCTATATACACACGTACACATAAA
TACACCCAGCACTGCAAGGCTAGAGGAAACTGGTGCACACTCTACAGTCTGACTGATTCA
TGTTCTGGAGAGCAGGACATAATGTATGATGAGAATGATCAAGGACTCTACACACTGGT
GGCTTGGAGAGGCCACTTCCCAGAATAATCCTGAGAGAAAAGGAATCATGGAGCAATGG
TGTTGAGTTCACTCAAGCCAAATGCCGTGCAGAGGGAAATGGCTTAGCGAGCTCTACAGT
AGGTGACCTGGAGGAAGGTACAGCCACACTGAAAATGGGATGTGCATGAACACGGAGGATC
CATGAACTACTGTAAAGTGTGACAGTGTGCAACTGCAGACAGCAGGTGAAATGTATGT
GTGCAATGCGACGAGAATGCAGAAGTCAGTAACATGTGCATGTTGTTGTGCTCCTTTTC
TGTTGGTAAAGTACAGAATTCAAGCAAATAAAAGGCCACCCCTGGCAAAAGCGGTAAAAAA
AAAAAA

FIGURE 142

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57033
<subunit 1 of 1, 311 aa, 1 stop
<MW: 35076, pI: 5.04, NX(S/T): 2
MQTFTMVLEEIWTSLFMWFFYALIPCLLTDEVAILPAPQNLSQLSTNMKHLLMWSPVIAPGE
TVYYSVYQGEYESLYTSHIWIPSSWCSLTEGPECVDVTDDITATVPYNLVRATLGSQTSAW
SILKHPFNRNSTILTRPGMEITKDGFLVIELEDLGPQFEFLVAYWRREPGAEHVKMVRSG
GIPVHLETMEPGAAYCVKAQTFVKAIGRYSAFSQTECVEVQGEAIPLVLALFAFVGFMLILV
VVPLFWWMGRLLQYSCCPVVVLPTDLKITNSPQLISCRREEVDACATAVMSPEELLRAWIS
```

Important features:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 230-255

N-glycosylation site.

amino acids 40-43 and 134-137

Tissue factor proteins.

amino acids 92-119

Integrins alpha chain proteins

amino acids 232-262

FIGURE 143

TCCTGCTGATGCACATCTGGTTGGCAAAGGAGGTTGCTCGAGCCGCCCTTAGCTT
CCTGGCCGGCTCTAGAACATTAGGCTCGCTGCGACTAGACCTCAGCTCCAACATATGCA
TTCTGAAGAAAGATGGCTGAGATGACAGAATGCTTATTTGGAAAGAAACAATGTTCTAGG
TCAAACGTGACTACCAAATGCAGACTTCACAATGGTTCTAGAAGAAATCTGGACAAGTCT
TTTCATGTGGTTTCTACGCATTGATTCCATGTTGCTCACAGATGAAGTGGCCATTCTGC
CTGCCCTCAGAACCTCTCTGTACTCTCAACCAACATGAAGCATCTTGTGGAGGCCA
GTGATCGCGCCTGGAGAACAGTGTACTATTCTGTCGAATACCAGGGGAGTACGAGAGCCT
GTACACGAGCCACATCTGGATCCCCAGCAGCTGGTGTCACTCACTGAAGGTCTGAGTGTG
ATGTCACTGATGACATCACGCCACTGTGCCATACAACCTTGTGTCAAGGCCACATTGGC
TCACAGACCTCAGCCTGGAGCATCCTGAAGCATCCCTTAATAGAAACTCAACCATCCTTAC
CCGACCTGGATGGAGATCACAAAGATGGCTNCACCTGGTTATTGAGCTGGAGGACCTGG
GGCCCCAGTTGAGTTCTTGTGGCTANTGGAGGGGGCGAACCCCTGCGCGCAAGGG
GTTNGCGAACCCCTTGCGGCCGCTGGGTATCTCTCGAGAAAAGAGAGAGGCCAATATGACCC
ACATACTCAATATGGACGAANTGCTATTGTCCACCTGTTGAGTGGCGCTGGTTGAT

FIGURE 144

CCACCGCGTCCGCCACCGTCCGAGGGACAAGAGAGAAGAGAGACTGAAACAGGGAGAAGA
GGCAGGAGAGGAGGAGGTGGGAGAGCACGAAGCTGGAGGCCACACTGAGGGAGGGCGGGA
GGAGGTGAAGAAGGAGAGAGGGAGAAGAGGGCAGGAGCTGAAAGGAGAGAGGGAGGAGGAG
GAGGAGATGCGGGATGGAGACCTGGAGTTAGGTGGCTTGGAGAGCTTAATGAAAAGAGAAC
GGAGAGGAGGTGTGGTTAGGAACCAAGAGGTAGCCCTGTGGCAGCAGAAGGCTGAGAGGA
GTAGGAAGATCAGGAGCTAGAGGGAGACTGGAGGGTTCCGGAAAAGAGCAGAGGAAGAGG
AAAGACACAGAGAGACGGGAGAGAGAAGAAGAGTGGGTTGAAGGGCGGATCTCAGTCCTG
GCTGCTTGGCATTGGGAACTGGACTCCCTGTGGGAGGAGAGGAAAGCTGGAAGTCCT
GGAGGGACAGGGTCCCAGAAGGAGGGACAGAGGAGCTGAGAGAGGGGGCAGGGCGTGGG
CAGGGTCCCTCGGAGGCCTCGGGATGGGGCTGCAGCTCGTCTGAGCCCCCCTCGAGC
GCTGGTACTCTGGCTGCACTGGGGCAGCAGCTCACATCGGACCAGCACCTGACCCCGAGG
ACTGGTGGAGCTACAAGGATAATCTCCAGGGAAACTCGTGCAGGGCTCCTTCTGGGC
CTGGTGAATGCAGCGTGGAGTCTGTGTGGAGGAGCAGAGCCCCGTGGATGTGGA
GCTGAAGAGGGTTCTTATGACCCCTTCTGCCCTTAAGGCTCAGCACTGGAGGAGAGA
AGCTCCGGGAACCTTGTACAACACCGGCCGACATGTCTCCTCCTGCCTGCACCCGACCT
GTGGTCAATGTGTCTGGAGGTCCCCTCCTTACAGCCACCGACTCAGTGAACTGCGGCTGCT
GTTTGGAGCTCGCGACGGAGGCCGCTCGAACATCAGATCAACCACCAAGGGCTTCTGCTG
AGGTGCAGCTCATTCACTCAACCAGGAACCTACGGGAATTCAGCGCTGCCTCCCGCGC
CCCAATGGCCTGCCATTCTCAGCCTTTGTCAACGTTGCCAGTACCTCTAACCCATTCT
CAGTCGCCTCTTAACCGCGACACCATCACTCGCATCTCCTACAAGAATGATGCCTACTTC
TTCAAGACCTGAGCCTGGAGCTCCTGTTCCCTGAATCCTCGGCTTCATCACCTATCAGGGC
TCTCTCAGCACCCGCCCTGCTCCGAGACTGTACCTGGATCCTCATTGACCGGGCCCTCAA
TATCACCTCCCTCAGATGCACCTCCCTGAGACTCCTGAGCCAGAATCCTCCATCTCAGATCT
TCCAGAGCCTCAGCGTAACAGCCGCCCTGCAGCCCTGGCCACAGGGCACTGAGGGC
AACAGGGACCCCGGCACCCCGAGAGGGCGCTGCCGAGGCCCAACTACCGCCTGCATGTGGA
TGGTGTCCCCCATGGTCGCTGAACTCCCCTCGAGGATTGCACCCGCCGCTTAAGCCTC
CCCACAAGGCGAGGGAGTTACCCCTAAAACAAAGCTATTAAAGGGACAGAACACTTA

FIGURE 145

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA34353
<subunit 1 of 1, 328 aa, 1 stop
<MW: 36238, pI: 9.90, NX(S/T): 3
MGAAARLSAPRALVLWAALGAAAHHGPAPDPEDWWWSYKDNLQGNFVPGPPFWGLVNAAWSLC
AVGKRQSPVVDVELKRVLYDPFLPPLRLSTGGEKLRGTLYNTGRHVSFLPAPRPVVNVSGGPL
LYSHRLSELRLLFGARDGAGSEHQINHQGFSAEVQLIHFNQELYGNFSAASRGPNGLAILSL
FVNVASTSNPFLSRLLNRDITRISYKNDAYFLQDLSLELLFPESFGFITYQGSLSTPPCSE
TVTWILIDRALNITSLQMHSRLLSQNPPSQIFQSLSGNSRPLQPLAHRALRGNRDPRHPER
RCRGPNYRLHVDGVPHGR
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 177-199

N-glycosylation site.

amino acids 118-121, 170-173 and 260-263

Eukaryotic-type carbonic anhydrases proteins

amino acids 222-270, 128-164 and 45-92

FIGURE 146

GGCGCCCTGGCTCGCGCTACTGGCTGTACGGAGCAGGAGCAAGAGGTGGCCAGGCCTCCGCCGCCAGCCTC
GTTCGTGTCCCCGGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGGCCACCCCTGGCAGACTAACGAA
GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTGGACCTTGCCTGCCATTCTCCAGGTTGAGGGAGC
CGCAGAGGGGGCTCGCGTATTCCCTGCAGTCAGCACCCACGTCGCCCGGACGCTCGGTGCTCAGGCCCTC
GGCAGGGGGCTCTCGTCTCGGGCCCTGTGAAGGCTCTGGGGCTGCGAGAGGCCGCGTCCGGTTGGCT
CACCTCTCCAGGAAACTCACACTGGAGAGCAGGAAAGAGCTGGAGAGCCTGTCTGGAGATTTCTGGGGAA
ATCCGTAGGTCAATTATGAAGTGTACCGCGGGAGTGCTCAGAGTAACCACAGTGTCTGGCATGGCTAGA
GCAATTCCAGCCATGGTGGTCCCAATGCCACTTTATTGGAGAAACTTTGGAAAATACATGGATGAGGATGGT
GAGTGGTGGATAGCAAACACGAGGGAAAGGGCATCACAGACAATGACATGCAGAGTATTGGACCTTCAT
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTTAATATGGAGTATGACATGGGATGTAGAGCTGGAAAGA
TCTGCAGAATCCTGGCTGAAGTTGCTTGTGGGACATGGACCTGCAAGCTGCTTCCATCAATTGGACAGAAT
TTGGGAGCACACTGGGAAGATAAGGCCCGACGTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTAGGTGTTCTGGCCCTGTATGTACACATTATACA
CAGGTCGTGTGGCAACTAGTAACAGAATCGGTTGCCCCATTAAATTGTGTCTAACATGAACATCTGGGGAG
ATATGCCCAAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAAGGGAAACTGGTGGGGCATGCCCTTACAAA
CATGGCGGCCCTGTTCTGCTTGTGCCCCACTAGTTTGAGGGGGCTGTAGAGAAAATCTGCTACAAGAAGGG
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC
CATGTCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGGCCACAGCAAATGTCCAAATTGTTCT
TGTGAAGTAAGATTAAGAGATCAGTGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCTTGT
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCTTGT
ATAATAGACAATGATGGTGGCTGGTAGATATCACTAGACAAGGAAGAAAGCATTATTCTCATCAAGTCCAATAGA
AATGGTATTCAAACAATTGCAAAATTCAGTCGTAATTCTTACAGCTCTAAAGTAACAGTTAGGTGCTG
ACTTGTGAAACAACAGCTCTGTCATTCTCATAAAGCCTGCTTCACATTGCCAAAGAGTATAACTGTCCT
CGTAACGTATGCAAGCAAATCCACATTATGCTCGTGTAAATTGAAACTCGAGTTATTCTGATCTGTCAGTATC
TGCAGAGCAGCAGTACATGCTGGAGTGGTGTGAAATCACGGGGTTATGTTGATGTAATGCCCTGTTGACAAAAGA
AAGACCTACATTGCTCTTCTCAGAATGGAATCTCTCAGAAAGTTACAGAATCTCCAGGAGGAAAGGCATT
AGAGTGGTTGCTGTTGTGTAAGACTGAATACTTGAAGAGGACATAAAAGACTATTCCAATGCAATTCTGA
ATTTTGATAAAACTGTAACATTACTGTACAGAGTACATCAACTATTCTCAGCCAAAAGGTGCCAAATGCATA
TAAATCTGATAAAACAAAGTATAAAATAAAACATGGGACATTAGCTTGGAAAAGTAATGAAAATATAATGG
TTTAGAAATCTGTTAAATATTGCTATATTCTTAGCAGTTATTCTACAGTTAAATTACATAGTCATGATT
GTTCTACGTTCTATATTATGGTGTCTTGTATATGCCACTAATAAAATGAAATCTAAACATGAAATGTAATG
GCCCTCAGAAAATCATCTAGTCATTAAAATATGACTCTAAACTGAAAGAAACCTTATCACATTCCCC
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTCCACTTAATAACTGTAAGTTTTTC
TGTTAATTAGGCATATAGAATATTAAATTCTGATATTGCACTTCTTATTATAAAATACTCTTAAATATC
CAAATGAAATCTGTTAAATGTTGATTCCTGGGAATGGCCTTAAAATAATGTAATAAGTCAGAGTGGTGGT
ATGAAAACATTCTAGTGTACATGTTAGGTTAACGATGGCAGAGCTTCTATGTACTGTTA
AAATTGAGGTACACATATTCTTTGTATCTGGAAATACTCTGCAGGCCAGGAAGTATAATGAAAAGTT
GAACAAAGATGAACTAATGTTAGTACATTACATTGCCACTGATTTTTAAATGTTAAATGACCTTGTATATAA
ATATTGCCATATCATGGTACTATAATGGTGTATATTGTTCTATGAAAATGTTGCTGGGTTACATATTAAATTTCTGCTGGTGG
TAAACATTAAAATTAATCATGTTCAAAAAAA

FIGURE 147

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417
<subunit 1 of 1, 500 aa, 1 stop
<MW: 56888, pI: 8.53, NX(S/T): 2
MKCTAREWLRVTTVLFMARAIPAMVVPNATLLEKLYMDEGEWWIAKQRGKRAITDNDM
QSILDLHNKLRSQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHW
GRYRPPTFHVQSWYDEVKDFSYEPYEHECNPYCPFRCSPGVCTHYTQVVWATSNRIGCAINLC
HNMMNIWGQIWPKAVYLVCNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRY
YPPREEETNEIERQQSQVHDTHVRTRSDDSSRNEVISAQQMSQIVSCEVRLRDQCKGTTCNR
YECPAGCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGI
QTIGKYQSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIG
TRVYSDLSSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSESLQNPPGGKAFRV
FAVV
```

Important features:

Signal peptide:

amino acids 1-20

Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

N-glycosylation site

amino acids 28-31

FIGURE 148

GC GGAGACAAGCGCAGAGCGCAGCGCACGCCACAGACAGCCCTGGCATCCACCGACGGCG
CAGCCGGAGCCAGCAGAGCGGAAGGCAGCCCCGGCAGAGAAAGCCGAGCAGAGCTGGGT
GGCGTCTCCGGGCCCGCTCCGACGGGCAGGCCCTCCCCATGTCCCTGCTCCCACGCCG
CGCCCTCCGGTCAGCATGAGGCTCCTGGCGCCGCTGCTCCTGCTGCTGGCGCTGT
ACACCGCGCGTGTGGACGGTCAAATGCAAGTGCTCCCGAAGGGACCCAAAGATCCGCTAC
AGCGACGTGAAGAAGCTGGAAATGAAGCCAAAGTACCCGACTGCGAGGAGAAGATGGTTAT
CATCACCACCAAGAGCGTGTCCAGGTACCGAGGTCAAGGAGCACTGCCTGCACCCCAAGCTGC
AGAGCACCAAGCGCTTCATCAAGTGGTACAACGCCGAGAAGCGCAGGGTCTACGAA
GAATAGGGTGAAAAACCTCAGAAGGGAAAACTCCAAACCAGTTGGAGACTTGTGCAAAGGA
CTTGCAGATTAaaAGCCTTTC
TTTCTCACAGGCATAAGACACAAATTATATATTGTTATGAAGCACTTTTACCAACGGTCAG
TTTTACATTTATAGCTGCGTGCAGGCTTCCAGATGGGAGACCCATCTCTCTTGCT
CCAGACTTCATCACAGGCTGCTTTATCAAAAGGGAAAACTCATGCCTTCCTTTAA
AAAATGCTTTTGATTTGTCCATACGTCACTATAACATCTGAGCTTATAAGCGCCGGGA
GGAACAATGAGCTGGTGGACACATTCAATTGCAGTGTGCTCCATTCTAGCTTGGGAAGC
TTCCGCTTAGAGGTCTGGCGCTCGGACAGCTGCCACGGCTCTCCTGGCTATGCCG
GTCACAGCCTCAGTGTGACTCCACAGTGGCCCTGTAGCCGGCAAGCAGGAGCAGGTCTCT
CTGCATCTGTTCTCTGAGGAACTCAAGTTGGTGCAGAAAATGTGCTTCAATTCCCCCT
GGTTAATTTACACACCCTAGGAAACATTCCAAGATCCTGTGATGGCGAGACAAATGATC
CTTAAAGAAGGTGTGGGTCTTCCAACCTGAGGATTCTGAAAGGTTCACAGGTTCAATA
TTTAATGCTTCAGAAGCATGTGAGGTTCCAACACTGTCAGCAAAACCTTAGGAGAAAACT
AAAAAATATATGAATAACATGCGCAATACACAGCTACAGACACACATTCTGTTGACAAGGGAA
AACCTCAAAACGTGTTCTTCCCTCACCAACAGAACATGCAGTACTAAAGCAATATAT
TTGTGATTCCCCATGTAATTCTCAATGTTAAACAGTGCAGTCCTCTCGAAAGCTAAGAT
GACCATGCGCCCTTCCCTGTACATATACCCTTAAGAACGCCCTCCACACACTGCC
CAGTATATGCCGCATTGTACTGCTGTGTTATGCTATGTACATGTCAGAAAACCATTAGCAT
TGCATGCAGGTTCATATTCTTCTAAGATGGAAAGTAATAAAATATTTGAAATGTAAAA
AAAAA

FIGURE 149

MSLLPERRAPPVSMRLLAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHLHPKLQSTKRFIKWYNNAWNEKRRVYEE

Signal sequence:

amino acids 1-34

MSLLPERRAPPVSMRLLAALLLLLLALYTARVDGSKCKCSRKGPKIRYSDVKLEMKPKYPH
CEEKMVIITTKSVSRYRGQEHLHPKLQSTKRFIKWYNNAWNEKRRVYEE

FIGURE 150

CCCCCAGGGACTGCTATGGCTTCCCTTGTGTTCACCCCGGCTGGCGCATGTTAAACTCCAATGTCCTCCTGTG
GTTAACTGCTCTGCCATCAAGTTCACCTCATTGACAGCCAGCACAGTATCCAGTTGTCACACACAAATTATGG
CAAATCCGGGGCTAAGAACACCGTTACCCAAATGAGATCTGGGTCCAGTGGAGCAGTACTTAGGGTCCCTA
TGCCTCACCCCCACTGGAGAGAGGCGTTCAGCCCCAGAACCCCGTCCTCTGGACTGGCATCCGAAATAC
TACTCAGTTGCTGCTGTGCCCCAGCACCTGGATGAGAGATCCTTACTGCATGACATGCTGCCATCTGGTT
TACCGCCAATTGGATACTTGTGACCTATGTTCAAGATCAAATGAAGACTGCCTTACTTAAACATCTACGT
GCCCACGGAAGATGGAGCCAACACAAAGAAAACGCAGATGATATAACGAGTAATGACCGTGGTAAGAGCCAAGA
TATTGATGATCAGAACAGTAAGAACGCCGTATGGTCTATATCCATGGGGATCTTACATGGAGGGCACCGCAA
CATGATTGACGGCAGCATTTGGCAAGCTACGGAAACGTCATCGTATCACCATAACTACCGTCTGGGAATACT
AGGGTTTTAAGTACCGGTGACCAGGCAGCAAAGGCAACTATGGGCTCTGGATCAGATTCAAGCACTGCGGTG
GATTGAGGAGAATGTGGAGCCTTGGGGGGACCCAAAGAGAGTGACCATCTTGCTCGGGCTGGGGCTGGGCTC
CTGTGTCAGCCTGTTGACCTGTCCCACACTCAGAAGGTCTTCCAGAAGGCCATCATTAGCAGGGCACCGC
CCTGTCAGCTGGCAGTGAACCTACAGCCGGCAAGTACACTCGGATATTGGCAGACAAGGTCGGCTGCAACAT
GCTGGACACCACGGACATGGTAGAATGCCCTGCGGAACAAGAACTACAAGGAGCTCATCCAGCAGACCATCACCC
GGCCACCTACCACATAGCCTTCGGGCGGTGATCGACGGCAGTCATCCCAGACGACCCCCAGATCCTGATGGA
GCAAGGGCAGTCTCAACTACGACATCATGCTGGCGTCAACCAAGGGAAAGGCTGAAGTTCTGTGGACGGCAT
CGTGGATAACGAGGACGGTGTGACGCCAACGACTTTGACTTCTCGTGTCCAACCTCGTGGACAACCTTACGG
CTACCCGAAAGGGAAAGACACTTGCGGGAGACTATCAAGTTCATGTACACAGACTGGCCGATAAGGAAAACCC
GGAGACGGCGGGAAAACCCCTGGTGGCTCTCTTACTGACCACAGTGGTGGCCCCCGCCGTGGCCGCGACCT
GCACGGCAGTACGGCTCCCCACCTACTTCTATGCCCTCATCAGTCCGCAAAAGCAGGAAATGAAGGCCAGCTG
GGCAGATTGGCCCATGGTGTGAGGTCCCCTATGCTTCGGCATCCCAGTATCGTCCCACCGAGCTCTCAG
TTGTAACCTTCAAGAACGACGTCATGCTCAGGCCGTGGTATGACCTACTGGACGAACTTCGCCAAAATGG
TGATCCAAATCAACCAAGCAGCTCATGCTCAGGATACCAAGTTCACTTACACAAAACCCAAACGGCTTGAAGAAGTGGCTG
GTCCAAGTATAATCCCAAAGACCAGCTCATGCTCAGGATACCAAGTGTGAGGAGATCACTACCGG
AACGAAAGTGGCTTCTGGTGGAACTCGTTCTCATTTGCACAACTTGAACGAGATATTCCAGTATGTTCAAC
AACACCAAAGGTTCTCCACCAAGACATGACATCATCCCTATGCCACCCGGGATCTCCGCCAAGATATGGCC
AACACCAAACGCCAGCAATCACTCCGCAACAATCCCAAACACTCTAAGGACCTCACAACACAGGGCTGA
GGACACAATGTCCTATTGAAACCAAACGAGATTATTCCACCGAATTAGTGTACCGATTGCCGTGGGGCGTC
GCTCCTTCTCCTCAACATCTAGCTTGTGGCGCTGTACTACAAAAGGACAAGAGGCCATGAGACTCACAG
GCGCCCCAGTCCCCAGAGAAACACCACAAATGATATCGTCACATCCAGAACGAAGAGATCATGCTCTGAGAT
GAAGCAGCTGGAACACGATCACGAGTGTGAGTCGCTGCAGGCACACGACACACTGAGGCTACCTGCCGCCAGA
CTACACCTCACGCTGCGCCGGTCGCCAGATGACATCCCACCTATGACGCCAACACCATCACCAGATTCCAAA
CACACTGACGGGGATGCAGCCTTGCACACTTTAACACCTTCAGTGGAGGACAAAACAGTACAAATTACCCCA
CGGACATTCCACCACTAGAGTATAGCTTGCCTATTCCCTCATCCCTGCCCTACCGCTCAGCAACAT
AGAACAGGGAAAGGAAG
GACTTAAGACAAAATGCAAAAGGCAGTCATCCCACCCGGCAGACCCCTTATGTTGGTGTGTTCCAGTATTAC
AAGATCAACTCTGACCCGTGAAATGTGAGAAGTACACATTCTGTTAAAATAACTGTTAAAGATCTTACCA
CTCCAATGTTAGTGTGATAGGACATCACCATTCAAGGCCCGGGTGTGTTCAACGTCATGGAAGCAGCT
GACACTCTGAAACTCAGCCAAGGACACTTGATATTTTAATTACAATGGAAGTTAAACATTCTTCTGTG
CACACAATGGATGGCTCTCTTAAGTGAAGAAAGAGTCATGAGATTGCCCCAGCACATGGAGCTGTAATCCAG
AGAGAAGGAAACGTAGAAATTATTATTAAGAACATGGACTGTGCAAGCAGAAATCTGTACGGTTCTGTGCAAAGAG
GTGTTTGCCAGCCTGAACATATTTAAGAGACTTTGT

40047082 - 402104

FIGURE 151

MLNSNVLLWLALAIKFTLIDSQAQYPVVNTNYKIRGLRPLPNEILGPVEQYLGVPYASP
PTGERRFQPPEPPSSWTGIRNTTQFAAVCPQHLDERSLLHDMLPIWFTANLDTLMTYVQDQN
EDCLYLNIVPTEGANTKKNADDITSNDRGEDEDIHDQNSKKPVMVYIHGGSYMEGTGNMI
DGSILASYGNVIVITINYRLGILGFLSTGDQAAKGNYGLLDQIQALRWIEENVGAFGGDPKR
VTIFGSGAGASCVSLLTLSHYSEGLFQKAIIQSGTALSSWAVNYQPAKYTRILADKVGCNML
DTTDMVECLRNKNYKELIQQTITPATYHIAFGPVIDGDVIPDDPQIILMEQGEFLNYDIMLGV
NQGEGLKFVDGIVDNEDGVTNPNDFSVSNFVDNLYGYPEGKDTLRETIKFMYTDWADKENP
ETRRKTLVALFTDHQWVAPAVAADLHAQYGSPTYFYAFYHHCQSEMPSWADSAHGDEVPYV
FGIPMIGPTELFSCNFSKNDVMLSAAVVMTYWTNFAKTGDPNQPVPQDTKFIHTKPNRFEEEVA
WSKYNPKDQLYLHIGLKPRVRDHYRATKVAFWLELVPHLHNLNEIFQYVSTTKVPPDMTS
FPYGTRRSPAKIWPTTKRPAITPANNPKHSKDPHKTGPEDTTVLIETKRDYSTELSVTIAVG
ASLLFLNILAFAALYYKKDKRRHETHRRPSPQRNTNDIAHIQNEEIMSLQMKQLEHDHECE
SLQAHDTLRLTCPDYTLTLLRRSPDDIPLMTPNTITMIPNTLTGMQPLHTFNTFSGGQNSTN
LPHGHSTTRV

Signal sequence:

amino acids 1-24

Transmembrane domains:

amino acids 189-204, 675-692

FIGURE 152

GGGAAAGATGGCCGGCACTCTGGGACCCCTGGGT~~CGT~~GGCAGCAGTGGCGGCATGTTGT
CGGCTCGGGATGGGT~~C~~CAGGATGTTACTCCTCTTGTGGGGTCTGGCAGGGCCA
CAGCAAGTCGGGGCGGGTCAAACGTTGAGTACTGAAACGGGAGCACTCGCTGTCGAAGCC
CTACCAGGGTGTGGGACAGGCAGTCCACTGTGGAATCTGATGGCAATGCCATGGTGA
TGACCCAGTATATCCGCCCTACCCAGATATGCAAAGTAACAGGGTGCCTGTGGAACCGG
GTGCCATGTTCCCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAATCCATGGACAAGGAA
GAAGAATCTGCATGGGATGGCTGGCAATCTGGTACACAAAGGATCGGATGCCAGGGC
CTGTGTTGGAAACATGGACAAATTGTGGGCTGGAGTATTGTAGACACCTACCCAA
GAGGAGAAGCAGCAAGAGCGGGTATTCCCTACATCTCAGGCATGGTAACAACGGCTCC
CAGCTATGATCATGAGCGGGATGGCGGCCTACAGAGCTGGGAGGCTGCACAGCATTGTCC
GCAATCTTCAATTACGACACCTCCCTGGT~~G~~ATTGCTACGTCAAGAGGCATTGACGATAATG
ATGGATATTGATGGCAAGCATGAGTGGAGGGACTGCATTGAAGTGCCC~~G~~GAGTCCGCC
CCGCGGCTACTACTTCGGCACCTCCATCACTGGGATCTCAGATAATCATGATGTCA
TTTCC~~T~~GAAGTTGTTGA~~A~~CTGACAGTGGAGAGAACCCAGAAGAGGAAAGCTCC
GATGTGTTCTGCCCTCAGTGACAATATGAAGCTGCC~~T~~GAGATGACAGCTCCACTGCC
CCTGAGTGGCCTGCCCTTCATCGTCTTTCTCC~~T~~GGT~~T~~TTCTGTATTGCCA
TAGTCATTGGTATCATACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTCTACTTGA
GCCCTCCTGCCACCACCTTGTGACTGTCACCCATGAGGT~~T~~GAAGGAGCAGGCACTG
GCCTGAGCATGCAGCCTGGAGAGTGTCTCTAGCAGCTGGTGGGACTATATTCTG
TCACTGGAGTTTGAATGCA~~GG~~ACCCCGATTCCATGGTGTGCATGGGACATCTAACT
CTGGTCTGGGAAGCCACCCACCCAGGGCAATGCTGCTGTGATGTGCC~~T~~TC~~T~~GCAGTCC
TTCCATGTGGGAGCAGAGGTGTGAAGAGAATTACGTGGTGTGATGCCAAACACAGAAC
AGAATTTCATAGCCCAGGCTGCC~~T~~GTTGACTCAGAAGGCC~~T~~CTACTTCAGTTG
AATCCACAAAGAATTAAAATGGTAACACCACAGGCTTCTGACCATCCATTGTTGGGTT
TTGCATTGACCCACCCCTGCC~~T~~ACCTGAGGAGCTTCTTGAAACCAGGATGGAAACT
TCTTCCCTGCC~~T~~ACCTCCTTC~~A~~TC~~C~~ATTGTCCTCTGTGTGCAACCTGAGCTG
GGAAAGGCATTGGATGCC~~T~~CTGTGGGGCTGGGCTGCAGAACACACCTGCGTT
TGGCCTTCATTAGGTGCC~~T~~CTAGGGAGATGGCTTCTGCTTGGATCA~~T~~GTCCCTAGCAT
GGGTCTGGGTCTATTGGCATGTCC~~T~~GGCC~~T~~CTCCAA~~T~~CAAGTCTTCAGGCC~~T~~CA~~G~~T
AAGTTGGCTAAAGGTGGTGTAAAATCAAGAGAAGCCTGGAAAGACATCATGGATGCCATG
GATTAGCTGTGCAACTGACCAGCTCCAGGTTGATCAAACCAAAAGCAACATTGT~~C~~ATGTG
GTCTGACC~~A~~TGTGGAGATGTTCTGGACTTGCTAGAGCCTGCTTAGCTGCATGTTGTAGT
TACGATTTGGAA~~T~~CCACTTGAGTGTGAAAGTGTAAAGGAAGCTTCTTACACCTT
GGGCTGGATATTGCCAGAGAAAGAAATTGGCTTTTTCTTA~~A~~GGACAAGAGACAGT
TGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCC~~T~~CATCATGTGCC~~T~~GGAAAGAGTT
CACTGTCATTGAGCAGCACAGCCTGAGTGTGCC~~T~~GTCAACCC~~T~~TATTCA~~T~~GTGCC~~T~~TA
TTTGACAAGGGGTTACATGCTGCTCACCTTACTGCC~~T~~GGGATTAAATCAGTTACAGGCCAG
AGTCTCC~~T~~GGAGGGC~~T~~GGAACTCTGAGTCC~~T~~GTGAA~~C~~CTGTAGCCTAA~~A~~ATGAAAT
TCTTAAAATCACCGATGGAACAAAAAAAGGGCGGCCGACTCTAGAGTCG
ACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCCATGG

FIGURE 153

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPLGSWQQWRRCLSARDGSRMLLLLLLGGSGQGPQQVGAGQTFEYLKREHSLSKPYQ
GVGTGSSSLWNLMGNAMVMTQYIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQGKKN
LHGDGLAIWYTKDRMQPGPVFGNMDKFVGLGVFVDTYPNEEKQQERVFPYISAMVNNGSLSY
DHERDGRPTELGGCTAIVRNLHYDTFLVIRYVKRHTIMMDIDGKHEWRDCIEVPGVRLPRG
YYFGTSSITGDLSDNHDVISLKLFELTVERTPEEEKLHRDVFLPSVDNMKLPEMTAPLPLS
GLALFLIVFFSLVFSVFAIVIGIILYNKWQEQRKRFY
```

Signal sequence:

amino acids 1-38

Transmembrane domain:

amino acids 310-329

FIGURE 154

CCGAGCCGGGCGCAGCGACGGAGCTGGGGCGGGCTGGGACCATGGCGTGAGTGCAATCTACGGATCAGTCT
CTGATGGTGGTCGTTAACCTCAGTGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGCTTCATTCAAGAAT
TGGGGTCTGGCTCAGAATTCTGCAGCTGGTGAAGAAAATCTGTTCTAGAAGAGGTTAATTAATGCCTGCAGTCT
GACATGTTCCGATTGAGGTGAAGAACATGAAGAGAAAATAGAATACTTAATAATGCTTTCCGCAACCGCTTCT
TGCTGCTGCTGGCCCTGGCTGGCCTGCTGGCTTGTGAGCCTCAGGCCAGCTGAGCAGGAGCCCTGATCCCCTGTGA
CGACTCTAAGAATGGAATGAGTAGCAAGAGTCAGAAGAGAATCATGCCAGTGTGGCGAGCGACATGGAAGGTATGCC
CAGACCCGTTATGAAGCTTTGTACTGCACATCCCAGTGTGGCGAGCAGGTACCCACTGTATGTCATTCCC
CGCATATTAAAGCTGGCTCAGTGATGTGTCATTGCCACGGAGACAGGTACCCACTGTATGTCATTCCC
AAACAAAGCGACAGAAATTGACTGCACCTGGTGGCTAACAGGAACCGTATCACCCAAAATGGAAGCTTCA
TTAGTCACATGTCAAAAGGATCCGGAGCCTTTGAAAGCCCCCTGAACCTCTTGCCCTTACCCAAATCACC
CATTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGAGCATTGAGAACGGTCAGCTGCTGAGGGATA
TCTATCTAAAGAAACACAAACTCCTGCCAATGATTGGCTGAGACAGCTCTATTAGAGACCACTGGGAAA
GCCGGACCTACAAAGTGGCTGGCCTGCTTATGGTTCTCCAGATTTGACTGGAAGAAGATTTATTCA
GGCACAGCCAAGTGCCTGCTCTGGAAGCTGCTATTGCCCGTAAGAAACCAAGTATCTGAAAAGGAGC
AGCGTCGTCACTCCTACGTTGAAAACAGCCAGCTGGAGAAGACCTACGGGAGATGCCAAGATCGTGG
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCATAGACTCCATGCTCTGCCACTCTGCCAACATGTCAGCT
TTCCTGTACCAAGAAATGGCTGTTGACATGGAGCACTCAAGGTAAATTAGACCCATCAGATGAGGATGAAA
GGGAAGACGGAGAAGAAATTGACTCGGGTATTCTCTCTGGTGCCACCCATCCTGAACCAAACCATCG
GCCGGATGCAGCGTGCACCGAGGGCAGGAAAGAAGAGCTTTGCCCTACTCTGCTCATGATGTCACCTGT
CACCGATTCTCAGTGCCTGGCCTTCAAGGCCAGGTTCCAAGGTTGCAGCAGGTTGATTTGAGCTTT
GGCAAGACAGAGAAAAGCCCAGTGAACATTCCGTCGGATTCTTACAATGGCTCGATGTCACATTCCACACCT
CTTCTGCCAAGACCACCAAGCGTCTCCAAGCCATGTGCCGTTGAAAACCTGGTCCGTTGTGAAAA
GGGACATGTTGAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGACATGTCACAGGGAAAGGATTCTAAAAGG
TATGCACTACAGCAGTATAAACATGCCAATACAGACCATAGGGAAAGGCTTCAAGGTTGCTGTTAC
TAAGGGTAGAAGATTATTGCTTTAAAGGCTAAATATTGTTGTGGAAACACAGATGGTTGGGTTGAACAGT
AAAGCACATTGCTGCAATGTTGACTGTGAATTGCTGGTACAAATGGCAGTTCAACAGAGGAATAGAAGGTACIT
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTCAGGACTGAAGACTGCAATCCAACTTGCAC
TCTTCTGCCCTGCCCATGTTACTATGATGGAAACCGCACACCTCAACCAAATTTTAACTTAGACATT
TTTACCTTGTCTTGTAAAGAATTCTGAGTGAATTCTAAATAAAGGTTGCCAAACTTTCTGTAAGG
GCCAGATTGTAATATTCAAGACTGTGAGGACAAAAGGCCACATACAGTCTCTGTCATAACTACTCAACTCTGT
TTCTGAAGCAGGAAAGCCACACAGACAGTACATAAAGGAATATGTTAGCTGGGTCCAGGCCAGACAAAACA
GATGGTGACAGACTGGCCCTGGGCTGTAGTTGCTGACCCCTCATCTAAAAAATAGGCTACTACAAATTGC
ACTTCCAGCACTTGAGAACAGAGTTGAATACCAAGAATTATTCAATGGTCCCTCCAGTAACTCTGCTAGAAACA
CAGAATTGGTCTGTATCTGACACTAGAACAAAATGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA
AACTGATTAGAAGAATACCTGATGTTATGATGATTGGTACAAGATAGTTAAGTATGTTCTAAATATTGT
CTGCTGAGTCTATTGCTGTATATGCTGAAATTGGTATGCCATTAGTATTTTATAGTTAGGAAAATATT
TTCTAAAGACCAGTTAGTGAATGACTCTTATTCTGTAGTAATATTCAATTGCTGTACCTGCTGGTGGTTAGAAG
GAGGCTAGAAGATGAATTAGGCACTTCTCCAAATAAAACTAATTATGGCTATTCCCTTGACAAGCTGAGA
ACTGGATTCAATTAAACCATTTCTCATCAGTTCAAATGGTAAATTCTGATTGATTTTAAATGCGTTTGG
AGAACTTTGCTATTAGGTAGTTACAGATCTTATAAGGTGTTTATATATTAGAAGCAATTATAATTACATCTG
TGATTCTGAACTAATGGTCTAATTCAAGAGAAATGGAAAGTGAAGTGAAGGATTCTGTTGTCATGGCATTCC
AACTTTCTCTTGTGTTTGTCCAGTGTGCAATTGAATATGTCATTCTATAAATAATTGAAATAA

FIGURE 155

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLALAALLAFLVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDPVTEPPVTDPVY
EALLYCNIPSVAERSMEGHAPHFKLVSVHFIRHGDPLYVIPKTKRPEIDCTLVANRKP
YHPKLEAFISHMSKGSGASFESPLNSLPLYPNHLCEMGELTQTGVVQHLQNGQLLRDIYLK
KHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDFDWKKIYFRHOPSALFCSGSCYCP
VRNQYLEKEQRRQYLLRLKNSOLEKTYGEMAKIVDVPTKQLRAANPIDSMLCHFCHNVSFPC
TRNGCVDMEHFVVIKTHQIEDERERREKKLYFGYSLLGAHPILNQTIGRMQRATEGRKEELF
ALYSAHDVTLSPVLSALGLSEARFPRFAARLIFELWQDREKPSEHSVRLYNGVDVTFTSF
CQDHHKRSPKPMCPLENLVRFVKRDMFVALGGSGTNYYDACHREGF
```

Signal sequence:

amino acids 1-18

FIGURE 156

AAAAAAGCTCACTAAAGTTCTATTAGAGCGAATACGGTAGATTCCATCCCCTTTGAAGAACAGTACTGTGGA
GCTATTAAAGAGATAAAAACGAAATATCCTTCTGGGAGTTCAAGATTGTCAGTAATTGGTAGGACTCTGAGC
GCCGCTGTTCACCAATCGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGACCGCCCTGAAGCACAAGCAGAT
AGCTAGGAATGAACCACCTGGGAGTATGTGAAACAACGGAGGAGCTCTGACTTCCAACGTCCCATTCTAT
GGCGAAGGAACGTCTCCTGACTTCAGTGGTTAACGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAATGAT
TCCTGCGGACTGCACCGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGACTCTGTGGGAGACCG
ATGCACCCAGATAACGCTATTCACTTCGGAGAGCTGGAGAAAGGCTCTAGGGTGGCGACATCTCCAGGGACCT
GGGGCTGGAGCCCCGGGAGCTCGGGAGCGGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTCGCCCT
GAATCCCGCAGCGGCAGTGGTACCGGCGGGCAGGATAGACCGGGAGGAGCTGTATGGGGCCATCAAGTG
TCAATTAAATCTAGACATTCTGATGGAGGATAAAAGTGAAAATAATGGAGTAGAAGTAGAAGTAAGGGACATTAA
CGACAATGCGCCTTACTTCGTGAAAGTGAATTAGAAATAAAATTAGTGAAAATGCAAGCCACTGAGATGCGGTT
CCCTCTACCCCACGCCTGGGATCCGGATATCGGGAGAACTCTGCAAGAGCTACGAGCTCAGCCGAACACTCA
CTTCTCCCTCATCGTCAAAATGGAGCCGACGGTAGTAAGTACCCGAATTGGTGTGAAACGCGCCCTGGACCG
CGAAGAAAAGGCTGCTACCCACCTGGTCTTACGGCCTCCGACGGGGCGACCGGGTGCACAGGCACCGCGCG
CATCCCGGTGATGGTTCTGGATGCGAACGACAACGACCGCAGCGTCTGCTCAGCCGAGTACCGCGAGCGTCTCC
GGAGAACTCTGGCCTTGGGACCGCAGCTGCTTGTAGTCAACGCTACCGACCCCTGACGAAGGAGTCAATGCGGAAGT
GAGGTATTCCCTCCGGTATGTGGACGACAAGGCGCCAAGTTTCAAACTAGATTGTAATTCAAGGACAATATC
AACAAATAGGGAGTTGGACCAACGAGGAGTCAGGATTCTACCAAGATGGAAGTGCAGCAAGCAATGGATAATGCAGGATA
TTCTGCGGAGCCAAAGTCTGATCACTGTTCTGGACGTGACGACAATGCCAGAAGTGGTCTCACCTCTCT
CGCCAGCTCGGTTCCGAAACTCTCCACAGGGACATTAATTGCCCTTTAAATGAAATGACCAAGATTCTGA
GGAAAACGGACAGGTGATCTGTTTATCCAAGGAAATCTGCCCTTAAATTAGAAAATCTACGGAAATTACTA
TAGTTTACTCACAGACATAGTCTTGGATAGGGAAACAGGTTCTAGCTACACATCACAGTGCACGCCACTGACCG
GGGAACCCCCTCATCCACGGAAACTCATATCTCGCTGAAACGTGGCAGACACCAACGACAAACCGCCGGTCTT
CCCTCAGGCCTCTTATCCCGTTATATCCACAGAGAACAAACATCCCAGAGGAGTTCCCTCGTCTGTGACCGCCCA
CGACCCGACTGTGAAAGAGAACGCCAGATCACTTATTCCCTGGCTGAGAACACCATCCAAGGGCAAGCCTATC
GTCCTACGTGTCCTCATCAACTCCGACACTGGGTACTGTATGCGCTGAGCTCTCGACTACGAGCAGTTCGAGA
CTTGCAACTGAAAGTGTGGCGGGACAACGGGACCCGCCCCCTAGCAGCAACGTGTCGTTGAGCCTGTTG
GCTGGGACAGACAATGCGCCCGAGATCCTGTACCCGCCCTCCCCACGGACGAGTTCCACTGGCGTGAGCT
GGCTCCCCGCTCCGACAGGCCGGTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGGCCAGAACGC
CTGGCTGTCCTACCGTCTGCTCAAGGCCAGCGAGCCGGACTCTCTGGTGGGTCTGCACACGGCGAGGTGCG
CACGGCGAGCCCTGCTGGACAGAGACCGCTCAAGCAGACCTCGTAGTGGCGTCCAGGACCAACGCCAGCC
CCCTCTCTCCGCCACTGTCACGCTCACCGTGGCGCAGACGATCCCCAAGTCCTGGGGACCTCGGCAG
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGACCTGGTGGTAGCGGTGGCCGGTCTCCTG
CGTCTTCTGGCCTTGTACCTGCTGGCGCTAGGCTGCGCCTGGCACAGTCACGCCCTGCTGCAGGC
TTCAGGAGGGCTTGTACAGGAGCGCCGGTCACTTGTGGCGTGGACGGGGTGCAGGCTTCTGAGAC
CTATTCCCACGAGGTTTCCCTCACCAACGGACTCGGGAGAGACTCACCTGATCTTCCCCAGCCAACATGAGA
CATGCTCGTCAGGCCAGGAGAGCTTGTAAAAAGCGAGCCCTTTGCTGTCAGGTGATTGCGTATTTCTAAAGA
CAGTCATGGGTTAATTGAGGTGAGTTATCAAATCTTCTTTTTTTAATTGCTCTGCTCCCAAGC
TGGAGTGCAGCGGTACGATCATAGCTCACTGCCGCTCAAACCTCTAGGCTCAAGCAATTATCCACCTTGCCT
CCGGTGTAAACAGGGACTACAGGTGCAAGCCACCTACTGTCGCTATCTATCTATCTATCTATCTATCT
CTATCTATCTATCTATCTATTACTTCTGTCAGACAGGGAGTCTCACGCCCTGAAATCCCAGTACTTGGGAGGC
CGAGGGGGTGGATCACCTGAGGTGGAGTTGAGACCAGCCCTGACCAACATGGAGAAACCCGTCTAATACTAA
AAAAATACAAAATAGCCGGCGTGGTGTGATGTCTGTAATCCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT
TGCTTAAACCTGGGAGGTGGAGGTTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGCAACAAGAGTG
AAACTCTATCTCA

FIGURE 157

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDLGLEPRELAER
GVRIIPRGRTQLFALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVR
DINDNAPYFRESELEIKISENAATEMRFPPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGA
DGSKYPELVLKRALDREEKAHHHLVLTASDGGDPVRTGATARIVMVLANDNAPAFQAQPEYR
ASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLCDNSGTISTIGELDHE
ESGFYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALLNVN
DQDSEENGQVICFIQGNLPFKLEKSYGNYSLVTIDIVLDREQVPSYNITVTATDRGTPPLST
ETHISLNVADTNNDNPPVFPQASYSAYIPENNPRGVSLVSVTAHDPCCEENAQITYSLAENTI
QGASLSSYVSINSDTGVLVALSSFDYEQFRDLQVKVMARDNGHPPSSNVSLFVLDQNDN
APEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQNAWLSSYRLLKASEPGLFSVG
LHTGEVRTARALLRDALKQSLVVAVQDHGQPPLSATVTLTAVADSIPOVLAQGSLESPA
NSETSDLTLYLVVAVAASCVFLAFVILLALRLRRWHKSRLLQASGGGLTGAPASHFVGVD
GVQAFLQTYSHESVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFSKDSHGL
IEVSLYQIFFLFFFNCVSQAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLS
IYLSIYLSIYLSIYLLSCTDGSLTPVIPVLWEAEAGGSPEVGSLRPA

Signal sequence:

amino acids 1-30

Transmembrane domains:

amino acids 693-711, 809-823, 869-888

FIGURE 158

CCCAGGGCTCTAGTGCAGGGAGGAAGGGAGGGAGCAGGGAGGTGGAGATTCCCAGTTAAAG
GCTCCAGAACATCGTGTACCAAGGCAGAGAACTGAAGTACTGGGGCCTCCTCCACTGGGTCCGAA
TCAGTAGGTGACCCGCCCTGGATTCTGGAAGACCTCACCATGGGACGCCCGACCTCGT
GCGGCCAAGACGTGGATGTTCTGCTCTGCTGGGGGGAGCCTGGCAGGACACTCCAGGGC
ACAGGAGGACAAGGTGCTGGGGGTATGAGTGCCAACCCATTGCAGCCTGGCAGGC
CCTTGTCCAGGGCAGCAACTACTCTGGCGGTGCCTGTAGGTGGCAACTGGTCCTT
ACAGCTGCCACTGTAAAAACCGAAATACACAGTACGCCTGGAGACCACAGCCTACAGAA
TAAAGATGGCCCAGAGCAAGAAATACCTGTGGTTAGTCCATCCCACACCCCTGCTACAACA
GCAGCGATGTGGAGGACCAACCATGATCTGATGCTTCTCAACTGCGTGACCAGGCATCC
CTGGGTCAAAGTGAAGCCCATCAGCCTGGCAGATCATTGCACCCAGCCTGGCAGAAGTG
CACCGTCTCAGGCTGGGCACTGTCACCAGTCCCCGAGAGAAATTTCTGACACTCTCAACT
GTGCAGAAGTAAAATCTTCCCCAGAAGAAGTGTGAGGATGCTTACCCGGGCAGATCACA
GATGGCATGGTCTGTGCAGGCAGCAGCAAAGGGCTGACACGTGCCAGGGCATTCTGGAGG
CCCCCTGGTGTGATGGTGCACTCCAGGGCATCACATCCTGGGCTCAGACCCCTGTGGGA
GGTCCGACAAACCTGGCGTCTATACCAACATCTGGCGTACCTGGACTGGATCAAGAAGATC
ATAGGCAGCAAGGGCTGATTCTAGGATAAGCACTAGATCTCCCTTAATAACTCACAACTCT
CTGGTTC

FIGURE 159

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48336
<subunit 1 of 1, 260 aa, 1 stop
<MW: 28048, pI: 7.87, NX(S/T): 1
MGRPRPRAAKTWMFLLLLGGAWAGHSRAQEDKVLGGHECOPHSQPWQAALFQGQQLLCGGVL
VGGNWVLTAAHCKKPKYTVRLGDHSLQNKGPEQEIPVVQSIHPHCYNSSDVEDHNHDLMLL
QLRDQASLGSKVKPISLADHCTQPGQKCTVSGWGTVTSPRENFPDTLNCAEVKIFPQKKCED
AYPGQITDGMVCAGSSKGADTCQGDGGPLVCDGALQGITSWGSDPCGRSDKPGVYTNICRY
LDWIKKIIIGSKG
```

Important Features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 51-71

N-glycosylation site.

amino acids 110-113

Serine proteases, trypsin family, histidine active site.

amino acids 69-74 and 207-217

Tyrosine kinase phosphorylation site.

amino acids 182-188

Kringle domain proteins motif

amino acids 205-217

FIGURE 160

GGCGCCGGTGCACCGGGCGGGCTGAGCGCCCTCTGCGGCCGCTGCGCGCCCCGGCCCGC
CGCGCCGCCAACGCCAACCCCGGCCGCCCCCTAGCCCCCGCCCGGGCCCGCGCCCG
GCCCGCGCCAGGTGAGCGCTCCGCCGCCGAGGCCCGCCGCCGCCGCCGCC
CCCCGGCCGGGGAAACCGGGCGGATTCCCTCGCGCGTCAAACCACTGATCCCATAAAAC
ATTCATCTCCGGCGCCCGCTGCGAGCGCCCCGCCAGTCCGCCGCCGCCCTCG
CCCTGTGCGCCCTGCGGCCCTGCGCACCCGCCGCCGAGGCCAGCCAGAGCGGGCGGAGC
GGAGCGCGCCGAGCCTCGTCCCGGCCGGGCCGGCCGTAGCGGCCGCGCCTGGA
TGC GGACCCGGCCGCCGGAGACGGGCCGCCGCCGAAACGACTTCAGTCCCCGACGCG
CCCGCCAACCCCTACGATGAAGAGGGCGTCCGCTGGAGGGAGCCGGCTGCTGGCATGGTG
CTGTGGCTGCAGGCCTGGCAGGTGGCAGCCCCATGCCAGGTGCCTGCGTATGCTACAATGA
GCCCAAGGTGACGACAAGCTGCCCGAGCAGGGCCTGCAGGCTGTGCCGTGGCATCCCTG
CTGCCAGCCAGCGCATCTCCTGCACGGAACCGCATCTCGCATGTGCCAGCTGCCAGCTTC
CGTGCCTGCCAACCTCACCATCCTGTGGCTGCACTCGAATGTGCTGGCCGAATTGATGC
GGCTGCCTTCACTGGCTGGCCCTCCTGGAGCAGCTGGACCTCAGCGATAATGCACAGCTCC
GGTCTGTGGACCCCTGCCACATTCCACGGCTGGGCCCTACACACGCTGCACCTGGACCGC
TGC GGCTGCAGGAGCTGGCCGGGGCTGTTCCGGCCCTGGCTGCCCTGCAGTACCTCTA
CCTGCAGGACAACCGCCTGCAGGCACTGCCGTGACACCTTCCGCGACCTGGCAACCTCA
CACACCTCTCCTGCACGGAACCGCATCTCCAGCGTGCCGAGCGGCCCTCCGTGGCTG
CACAGCCTCGACCGTCTCCTACTGCACCAAGAACCGCGTGGCCCATGTGCACCCGATGCC
CCGTGACCTTGGCCGCTCATGACACTCTATCTGTTGCCAACAAATCTATCAGCGCTGCC
CTGAGGCCCTGGCCCCCTGCGTGCCTGCAGTACCTGAGGCTCAACGACAACCCCTGGGT
TGTGACTGCCGGCACGCCACTCTGGCCTGGCTGCAGAAGTTCCGGCTCCCTCC
GGTGCCTGCAGCCTCCGCAACGCCCTGGCTGGCGTACCTCAAACGCCCTAGCTGCC
ACCTGCAGGGCTGCGCTGTGGCACCGGCCCTTACCATCCCATCTGGACCGGCAGGGC
GATGAGGAGCCGCTGGGGCTTCCAAGTGCTGCCAGCCAGATGCCGTGACAAGGC
ACTGGAGCCTGGAAGACCAGCTCGGCAGGCAATGCGCTGAAGGGACGCCGTGCC
ACAGCCCGCCGGCAACGGCTCTGGCCACGGCACATCAATGACTCACCC
CCTGGCTCTGCTGAGCCCCGCTCACTGCAGTGCCTGGCCGAGGGCTCC
CCCCACCTGGGCCCTGCCGGAGGCCAGGCTGTTCACGCAAGAACCGC
GCCGTCTGGCCAGGCAGGCAGCGGGGTGGCGGGACTGGTACTCAGAAGG
CTACCCAGCCTCACCTGCAGCCTACCCCCCTGGCCTGGCGTGGTCTG
TGGGCCCTGCTGACCCCCAGCGGACACAAGAGCGTGCAGCAGCCAGGTGTGT
GGGTCTCTCTCCACGCCGCAAGCCAGGCCGGCGGCCGACCGTGGGCAGGCC
GTCCTCCCTGATGGACGCCCTGCCGCCACCCCATCTCCACCC
TTCGGCGGCAGCGTTGTTCCAGAACGCCCTCCCACCC
GCATTTATTTACTTGTGTAAAAATATCGGACGACGTGGAATAAGAG
CTTTCTTAA
AAAA

FIGURE 161

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44184
><subunit 1 of 1, 473 aa, 1 stop
><MW: 50708, pI: 9.28, NX(S/T): 6
MKRASAGGSRLLAWVLWLQAWQVAAPCPGACVCYNEPKVTTSCPQQGLQAVPGIPAASQRI
FLHGNRISHVPAASFRACRNLTILWLHSNVLARIDAAFTGLALLEQLDLSDNAQLRSVDPA
TFHGLGRLHTLHLDRCGLQELGPGLFRGLAALQYLYLQDNALQALPDDTFRDLGNLTHLFLH
GNRISSVPERAFRGLHSLDRLLLHQNRVAHVPHAFRDLGRLMTLYLFANNLSALPTEALAP
LRALQYLRNDNPWVCDRARPLWAWLQKFRGSSSEVPCSLPQRLAGRDLKRLAANDLQGCA
VATGPYHPIWTGRATDEEPLGLPKCCQPDAADKASVLEPGRPASAGNALKGRVPPGDSPPGN
GSGPRHINDSPFGTLPGSAEPPPLTAVRPEGSEPPGFPTSGPRRPGCSRKNRTRSHCRLGQA
GSGGGGTGDSESGALPSLTCSLTPGLALVLWTVLGPC
```

Important features:

Signal peptide:

amino acids 1-26

Leucine zipper pattern.

amino acids 135-156

Glycosaminoglycan attachment site.

amino acids 436-439

N-glycosylation site.

amino acids 82-85, 179-183, 237-240, 372-375 and 423-426

VWFC domain

amino acids 411-425

FIGURE 162

GGAAGTCCACGGGGAGCTGGATGCCAAGGGAGGACGGCTGGGTCTCTGGAGAGGACTAC
TCACTGGCATATTCTGAGGTATCTGTAGAATAACCACAGCCTCAGATACTGGGGACTTAC
AGTCCCACAGAACCGTCCCTCCAGGAAGCTGAATCCAGCAAGAACAAATGGAGGCCAGCGGA
AGCTCATTGAGACAAAGGCAAGTCTTTCTCCTTTGGCTTATCTCTGGCG
GGCGCGCGGAACCTAGAAGCTATTCTGTGGTGGAGGAAACTGAGGGCAGCTCCTTGTAC
CAATTAGCAAAGGACCTGGGCTGGAGCAGAGGAAATTCTCCAGGGGGGGTTAGGGTTG
TTTCAGAGGAAACAAACTACATTGAGCTCAATCAGGAGACCGCGGATTGCTAAAT
GAGAAATTGGACCGTGAGGATCTGTGCGGTACACAGAGCCTGTGCTACGTTCCAAGT
GTTGCTAGAGAGTCCCTCGAGTTTCAAGCTGAGCTGCAAGTAATAGACATAAACGACC
ACTCTCCAGTATTCTGGACAAACAAATGTTGGTGAAGTATCAGAGAGCAGTCCTCTGGG
ACTACGTTCTCTGAAGAATGCCGAAGACTTAGATGTAGGCCAAACAAATTGAGAACTA
TATAATCAGCCCCAACCTCTATTTCGGGTCTCACCCGCAAACGCAGTGATGGCAGGAAAT
ACCCAGAGCTGGTGGACAAAGCGCTGGACCGAGAGGAAGAGCTGAGCTCAGGTTAAC
CTCACAGCACTGGATGGCTCTCCGCCAGATCTGGCACTGCTCAGGTCTACATCGAAGT
CCTGGATGTCAACGATAATGCCCTGAATTGAGCAGCCTTCTATAGAGTGAGATCTCTG
AGGACAGTCCGGTAGGCTCCTGGTTGTGAAGGTTCTGCCCCAGATCTGGATGTAGACACAGGAGTC
AACGGAGAGATTCCATTCACTTTCCAAGCTTCAGAAGAGATTGGCAAAACCTTAAGAT
CAATCCCTGACAGGAGAAATTGAACTAAAAAAACACTCGATTGCAAAACTTCAGTCCT
ATGAAGTCAATATTGAGGCAAGAGATGCTGGAACCTTTCTGAAAATGCACCGTTCTGATT
CAAGTGTAGATGTGAACGACCATGCCAGAAGTTACCATGTCTGCATTTACAGGCCAAT
ACCTGAGAACCGCGCTGAAACTGTGGTGCACTTTCAGTGTTCAGATCTGATTAGGAG
AAAATGGAAAATTAGTGCTCATTAGGAGGATCTACCCCTCCTCTGAAATCCGCGGAA
AACTTTACACCTACTAACGGAGAGACCACTAGACAGAGAAAGCAGAGCGGAATACACAT
CACTACACTGTCACTGACTTGGGACCCCTATGCTGATAACACAGCTCAATATGACCGTG
TGATGCCGATGTCAATGACAACGCTCCCGCCTCACCCAAACCTCCTACACCCCTGTTGTC
CGCGAGAACACAGCCCCGCCCTGCACATCCGAGCGTCAGCGCTACAGACAGAGACTCAGG
CACCAACGCCAGGTCAACCTACTCGCTGCTGCCGCCAGGACCCGACCTGCCCTCACAT
CCCTGGTCTCCATCAACCGGACAACGCCACCTGTTGCCCTCAGGTCTGGACTACGAG
GCCCTGCAGGGGTTCCAGTCCCGCTGGCGCTTCAGACCACGGCTCCCCGGCGCTGAGCAG
CGAGGGCGCTGGTGCCTGGTGGCTGGACGCCAACGACAACCTGCCCTCGTGTGTACCG
CGCTGCAGAACGGCTCCCGGCCCTGCACCGAGCTGGTCCCCGGCGGCCAGGCCGGCTAC
CTGGTACCAAGGTGGTGGCGGTGGACGGCGACTCGGGCAGAACGCCCTGGCTGTCGTACCA
GCTGCTCAAGGCCACGGAGCTGGTCTGTTGGCGTGGCGCACAATGGCGAGGTGCGCA
CCGCCAGGCTGCTGAGCGAGCGACGCCAACGCCACGCTGACGTGCTCTGGTGGACGGCTTCTC
AATGGCGAGCCTCCCGCCTGGCCACCGGCCACGCCAGGCCACCTGCTCAGGAAATTGCTCACCG
CCAGCCCTACCTGCCTCTCCGGAGGGGCCACCCAGGCCACCTGCTCAGGAAATTGCTCACCG
TCTACCTGGTGGTGGCGTTGGCCTCGGTGTCTCGCTCTCCCTCTTGGTGTGCTCTGGT
GTGGCGGTGCCTGGTGTAGGAGGAGCAGGGCGGCCCTGGTGGTGTGCTGGTGGCG
GGGCCCTCCAGGGCATCTTGTGGACATGAGCGGCCACCAAGGCCACCTATCCAGAGCTACC
AGTATGAGGTGTGTGGCAGGAGGCTCAGGGACCAATGAGTTCAAGTTCTGAAGCCGATT
ATCCCCAACCTCCCTCCCCAGTGCCCTGGGAAAGAAATACAAGGAAATTCTACCTCCCCAA
TAACCTGGGTTCAATATTCACTGACATGTTGACTTTACATTCCATAGGTATTTATT
TGTGGCATTCCATGCCAATGTTATTCCCCAATTGTTGTGTTGAAATTGTACGGAT
TTACTCTGATTTCTCATGTTCTCCCTTGGTTAAAGTGAACATTACCTTATT
CCTGGTTCTT

40047092 402104

FIGURE 163

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48314
<subunit 1 of 1, 798 aa, 1 stop
<MW: 87552, pI: 4.84, NX(S/T): 5
MEASGKLICRQRQVLFSFLLGLSLAGAAEPRSYSVVEETEGSSFVTNLAKDLGLEQREFSR
RGVRVVSRGNKLHLQLNQETADLLLNEKLDREDLCGTEPCVLRFQVLLESPFEFFQAELOV
IDINDHSPVFLDKQMLVKVSESSPPGTTFPLKNAEDLDVGQNNIENYIISPNSYFRVLTRKR
SDGRKYPELVLVDKALDREEEAELRLTLDGGSPRSGTAQVYIEVLDVNDNAPEFEQPFY
RVQISEDSPVGFLVVKVSATDVTGVNGEISYSLFQASEEIGKTFKINPLTGEIELKKQLDF
EKLQSYEVNIEARDAGTFSGKCTVLIQVIDVNDHAPEVTMSAFTSPIPENAPETVVALFSVS
DLDSENGKISCSIQEDLPFLLKSAENFYTLLTERPLDRESRAEYNITITVTDLGTPMLITQ
LNMTVLIADVNDNAPAFQTQSYTLFVRENNSPALHIRSVSATDRDSGTNAQVTYSLLPPQDP
HLPLTSLVSINADNGHLFALRSLDYEALQGFQFRVGASDHGSPALSSEALVRVVVLDANDNS
PFVLYPLQNGSAPCTELVPRAAEPGYLVTKVVAVDGDSGQNAWLQYQLLKATELGLFGVWAH
NGEVRTARLLSERDAAKHRLVVLVKDNGEPPRSATATLHVLLVDGFSQPYLPLPEAAPTQAQ
ADLLTVYLVVALASVSSLFLFSVLLFVAVRLCRRSRAASVGRCLVPEGPLPGHLVDMMSGRT
LSQSYQYEVCLAGGSGTNEFKFLKPIIPNFPPQCPGKEIQGNSTFPNNFGFNIQ
```

Important features:

Signal peptide:

amino acids 1-26

Transmembrane domain:

amino acids 685-712

Cadherins extracellular repeated domain signature.

amino acids 122-132, 231-241, 336-346, 439-449 and 549-559

ATP/GTP-binding site motif A (P-loop).

amino acids 285-292

N-glycosylation site.

amino acids 418-421, 436-439, 567-570 and 786-789

FIGURE 164

ACCCACCGTCCGCCACCGTCCGCCACCGTCCGCCACCGTCCGCCGTAGCCGTGC
GCCGATTGCCTCTCGGCCTGGCAATGGTCCC GGCTGCCGGTCGACGACGCCCGCGTCAT
GC GGCTCCTCGGCTGGCAAGTATTGCTGTGGGTGCTGGGACTTCCC GTCCGCCGGTGG
AGGTTGCAGAGGAAAGTGGTCGCTTATGGTCAGAGGAGCAGCCTGCTCACCCCTCTCCAGGTG
GGGGCTGTGTACCTGGGTGAGGAGGAGCCTGCATGACCCGATGGGCCAGGACAGGGCAGC
AGAAGAGGCCAATGCGGTGTGGGCTGGACACCCAAGGCGATCACATGGTGTGCTGTCTG
TGATT CCTGGGAAGCTGAGGACAAAGTGAGTTAGCAGGCCTAGCGCGTCACCTGTGGTGT
GGAGGAGCGGAGGACTCAAGGTGCAACGTCCGAGAGAGCCTTTCTCTGGATGGCGCTGG
AGCACACTCCCTGACAGAGAAGAGGAGTATTACACAGAGCCAGAAGTGGCGGAATCTGACG
CAGCCCCGACAGAGGACTCCAATAACACTGAAAGTCTGAAATCCCCAAAGGTGAACGTGAG
GAGAGAAACATTACAGGATTAGAAAATTCACTCTGAAAATTTAAATATGTCACAGGACCT
TATGGATTTCTGAACCCAAACGGTAGTGACTGTACTCTAGCCTGTTTACACCCCGTGGT
GCCGCTTTCTGCCAGTTGGCCCTCACTTTAACTCTCTGCCCGGGCATTCCAGCTCTT
CACTTTTGGA CACTGGATGCATCTCAGCACAGCAGCCTTCTACCAGGTTGGCACCGTAGC
TGTTCTAATATTTATTATTCAGGAGCTAAACCAATGCCAGATTAAATCATACAGATC
GAACACTGGAAACACTGAAAATCTCATTAAATCAGACAGGTATAGAAGCCAAGAAGAAT
GTGGTGGTAACTCAAGCCGACCAAATAGGCCCTCTCCCAGCAGCTTGATAAAAAGTGTGGA
CTGGTTGCTGTATTTCTTATTCTTTAATTAGTTTATTATGTATGCTACCATTGAA
CTGAGAGTATTGGCTAATTCCAGGACAAGAGCAGGAACATGTGGAGTAGTGATGGTCT
GAAAGAAGTGGAAAGAGGAACCTCAATCCTCGTTCAAGAAATTAGTGCTACAGTTCTATA
CATTTCTCCAGTGACGTGTTGACTTGAAACTTCAGGCAGATTAAAGAATCATTGTTGAA
CAACTGAATGTATAAAAAAATTATAAACTGGTGTAACTAGTATTGCAATAAGCAAATGC
AAAAATATTCAATAG

404267-2007-102107

FIGURE 165

```
>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48333
><subunit 1 of 1, 360 aa, 1 stop
><MW: 39885, pI: 4.79, NX(S/T): 7
MVPAAGRRPPRVMRLLGWWQVLLWVLGLPVRGVEVAEESGRLWSEEQPAHPLQVGAVYLGE
ELLHDPMGQDRAAEEANAVGLDTQGDHMMVMLSVIPGEAEDKVSSEPSGVTAGGAEEDSRC
NVRESLFSLDGAGAHFPDREEEYYTEPEVAESDAAPTEDSNNTESLKSPKVNCERNITGLE
NFTLKILNMSQDLMDFLNPNGSDCTLVLFYTPWCRFSASLAPHFNSLPRAFPALHFLALDAS
QHSSLSTRFGTVAVPNILLFQGAKPMARFNHTDRTLETLKIFIFNQTGIEAKKNVVVTQADQ
IGPLPSTLIKSVDWLLVFSLFFLISFIMYATIRTESIRWLIPGQEQEHEVE
```

Important features:

Signal peptide:

amino acids 1-25

Transmembrane domain:

amino acids 321-340

Homologous region to dilsufide isomerase

amino acids 212-302

N-glycosylation site.

amino acids 165-168, 181-184, 187-190, 194-197, 206-209, 278-281
and 293-296

Thioredoxin domain

amino acids 211-227

FIGURE 166

CCCGGCTCCGCTCCCTCTGCCCTCGGGTGCAGCGCCACGATGCTGCAGGGCCCTGGCT
CGCTGCTGCTCTCCTCGCCTCGCACTGCTGCCCTGGCTCGCGCGGGCTCTCCTC
TTTGGCCAGCCGACTCTCCTACAAGGCCAGCAATTGCAAGCCCATCCGGTCAACCTGCA
GCTGTGCCACGGCATCGAATACCAAGAACATGCGGCTGCCAACCTGCTGGCCACGAGACCA
TGAAGGAGGTGCTGGAGCAGGCCGGCGCTGGATCCGCTGGTATGAAGCAGTGCCACCCG
GACACCAAGAACTTCGCTCGCTCTGCCCGTCTGCCTCGATGACCTAGACGAGAC
CATCCAGCCATGCCACTCGCTCGTGCAGGTGAAGGACCGCTGCGCCCGGTATGTCCG
CCTTCGGCTTCCCCTGGCCCGACATGCTTGAGTGCACCGTTCCCCCAGGACAACGACCTT
TGCATCCCCCTCGCTAGCAGCGACCACCTCCTGCCAGCCACCGAGGAAGCTCAAAGGTATG
TGAAGCCTGAAAAATAAAATGATGATGACAAACGACATAATGAAACGCTTGAAAAATG
ATTTGCACTGAAAATAAAAGTGAAGGAGATAACCTACATCAACCGAGATACCAAATCATC
CTGGAGACCAAGAGCAAGACCATTACAAGCTGAACGGTGTCCGAAAGGGACCTGAAGAA
ATCGGTGCTGTGGCTCAAAGACAGCTTGCAGTGACCTGTGAGGAGATGAACGACATCAACG
CGCCCTATCTGGTCAAGGACAGAAACAGGGTGGGAGCTGGTATCACCTCGGTGAAGCGG
TGGCAGAAGGGCAGAGAGAGTTCAAGCGCATCTCCCGCAGCATCCGAAGCTGCAGTGCTA
GTCCCGGCATCCTGATGGCTCCGACAGGCCTGCTCCAGAGCACGGCTGACCATTCTGCTCC
GGGATCTCAGCTCCGTTCCCAAGCACACTCCTAGCTGCTCCAGTCTCAGCCTGGCAGCT
TCCCCCTGCCTTTGCACGTTGCATCCCCAGCATTCTGAGTTATAAGGCCACAGGAGTG
GATAGCTGTTTACCTAAAGGAAAAGCCCACCCGAATCTTGTAGAAATATTCAAACATAATA
AAATCATGAATATTTAA

FIGURE 167

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50920
><subunit 1 of 1, 295 aa, 1 stop
><MW: 33518, pI: 7.74, NX(S/T): 0
MLQPGSLLLLFLASHCCLGSARGLFLFGQPDFSYKRSNCKPIPVNLQLCHGIEYQNMRLPN
LLGHETMKEVLEQAGAWIPLVMKQCHPDTKKFLCSLFAPVCLDDLDETIQPCHSLCVQVKDR
CAPVMSAFGFPWPDMLECDRFPQDNDLCIPLASSDHLLPATEEAPKVEACKNKNDDNDIM
ETLCKNDFALKIKVKEITYINRTKIILETKSktiyKLNGVSERDLKKSVLWLKDSDLQCTCE
EMNDINAPYLVMGQKQGGELVITSVKRWQKGQREFKRISRSIRKLQC
```

Important features:

Signal peptide:

amino acids 1-20

Cysteine rich domain, homologous to frizzled N terminus

amino acids 6-153

FIGURE 168

GTGGAGGCCGCGACG**A**TGGCGGGGCCACGGAGGCCGAGACGGGTTGGCGAGCCCCGGG
CCCTGTGCGCGACGGGGCCACCGCACCTACGCGCGCGCTGGGTGTTCTGCTCGCGATC
AGCCTGCTCAACTGCTCCAACGCCACGCTGTGGCTCAGCTTGCACCTGTGGCTGACGTCA
TGCTGAGGACTTGGTCCTGTCCATGGAGCAGATCAACTGGCTGTCACTGGTCTACCTCGTGG
TATCCACCCATTGGCGTGGCGGCCATCTGGATCCTGGACTCCGTGGGCTCCGTGCGCG
ACCATCCTGGGTGCGTGGCTGAACTTGCCGGAGTGTGCTACGCATGGGCCCTGCATGGT
TGTTGGGACCCAAAACCCATTGCCTCCTCATGGGTGGCCAGAGCCTCTGTGCCCTGCC
AGAGCCTGGTCATCTCTCTCCAGCCAAGCTGGCTGCCCTGTGGTCTCCAGAGCACCAGCGA
GCCACGGCCAACATGCTGCCACCATGTCGAACCCCTCTGGCGTCCTGTGGCCAATGTGCT
GTCCCCCTGTGCTGGTCAAGAAGGGTGAGGACATTCCGTTAATGCTCGGTGTCTATACCATCC
CTGCTGGCGTCGTCTGCCCTGCTGTCCACCATCTGCCCTGTGGGAGAGTGTGCCCTTCC
CCCTCTGCCGGGCTGCCAGCTCACCTCAGAGAAGTCCCTGGATGGCTCAAGCTGCAGCT
CATGTGGAACAAGGCCTATGTCATCCTGGCTGTGCTGGGGGAATGATCGGATCTCTG
CCAGCTTCTCAGCCCTCCTGGAGCAGATCCTCTGTGCAAGCGGCCACTCCAGTGGTTTCC
GGCCTCTGTGGCGCTCTTCATCACGTTGGATCCTGGGGCACTGGCTCTCGGCCCTA
TGTGGACCGGACCAAGCACTTCACTGAGGCCACCAAGATTGGCCTGTGCCCTCTCTGG
CCTGCGTGCCCTTGCCCTGGTGTCCCAGCTGCAGGGACAGACCCCTGCCCTGGCTGCCACC
TGCTCGCTGCTCGGGCTGTTGGCTCTCGTGGGGCCATGGAGTTGGCGGTGA
GTGTTCTTCCCCGTGGGGAGGGGGCTGCCACAGGCATGATCTTGTGCTGGGCAGGCCG
AGGAAATACTCATCATGCTGGCAATGACGGCACTGACTGTGCGACGCTGGAGCCGCTTG
TCCACCTGCCAGCAGGGGGAGGATCCACTTGACTGGACAGTGTCTCTGCTGATGGCCGG
CCTGTGCACCTTCTCAGCTGCATCCTGGCGTCTTCTCACACCCATACCGGCCCTGC
AGGCCGAGTCTGGGAGCCCCCTCCACCGTAACGCCGTGGCGGCCAGACTCAGGCCG
GGTGTGGACCGAGGGGGAGCAGGAAGGGCTGGGTCTGGGGCCAGCACGGCACTCCGGA
GTGCACGGCGAGGGGGCCTCGCTAGAGGACCCCAGAGGGCCCGGGAGCCCCACCCAGCCT
GCCACCGAGCGACTCCCCGTGCGCAAGGCCAGCAGCCACCGACGCCCTCCGCCCG
AGACTCGCAGGCAGGGTCCAAGCGTCCAGGTTATTGACCCGGCTGGGTCTCACTCCTCCT
CTCCTCCCCGTGGGTGATCACGT**A**GCTGAGCGCCTGTAGTCCAGGTTGCCGCCACATCGA
TGGAGGCGAACTGGAACATCTGGTCCACCTGCCGGGGGGCAAAGGGCTCCTTGGGGCT
CCGGGAGCGAATTACAAGCGCGCACCTGAAAA

FIGURE 169

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50988
><subunit 1 of 1, 560 aa, 1 stop
><MW: 58427, pI: 6.86, NX(S/T): 2
MAGPTEAETGLAEPRALCAQRGHRTYARRWVFLLAISLLNCSNATLWLSFAPVADVIADLV
LSMEQINWLSLVYLVVSTPFGVAAIWILDVGRLRAATILGAWLNFAGSVLRMVPVPCMVGQTQN
PFAFLMGGQSLCALAQSLVIFSPAKLAALWFPEHQRTANMLATMSNPLGVIVANVLSPVLV
KKGEDIPLMLGVYTIPAGVVCLLSTICLWESVPPTPPSAGAASSTSEKFLDGLKLQLMWNK
YVILAVCLGGMIGISASFSALEQILCASGHSSGFSGLCGALFITFGILGALALGPYVDRTK
HFTEATKIGLCLFSLACVPFALVSQQLQGQTLALAATCSLLGLFGFSVGPVAMELAVECSFPV
GEGAATGMIFVLCQAEGILIMLAMTALTAVRSEPSLSTCQQGEDPLDWTVSLLL MAGLCTFF
SCILAVFFHTPYRRLQAESGEPPSTRNAVGGADSGPGVDRGGAGRAGVLGPSTATPECTARG
ASLEDPRPGPGSPHPACHRATPRAQPAATTDAPSRPGRLAGRVQASFIDPAGSHSSFSSPWVIT

Important features:

Signal peptide:

amino acids 1-44

Transmembrane domains:

amino acids 61-79, 98-112, 126-146, 169-182, 201-215, 248-268,
280-300, 318-337, 341-357, 375-387, 420-441

N-glycosylation site.

amino acids 40-43 and 43-46

Glycosaminoglycan attachment site.

amino acids 468-471

FIGURE 170

FIGURE 171

Page 280 of 280

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48331
<subunit 1 of 1, 1184 aa, 1 stop
<MW: 129022, pI: 5.20, NX(S/T): 5
MMQLLQLLLGLLGPGGYLFLLGDCQEVTTLTVKYQVSEEVPSGTIVGKLSQELGREERRQA
GAAFQVLQLPQALPIQVDSEEGLLSTGRRLDREQLCRQWDPCLVSFDVLATGDLALIHEIQ
VLDINDHQPRFPKGQELEISESASLRTRIPLDRALDPDTGPNTLHTYTLSPSEHFALDVIV
GPDETKHAELIVVKELDREIHSFFDLVLTAYDNGNPPKGTSLVKVNVLDSNDNSPAFAESS
LALEIQEDAAPGTLLIKLTATDPDQGPNGEVEFFLSKHMPEVLDTSIDAKTGQVILRRPL
DYEKNPAYEVDVQARDLGPNPPIPAHCKVLIKVLVDVNNDNIPSIHVTWASQPSLVSEALPKDSF
IALVMADDLDSGHNGLVHCWLSQELGHFRLKRTNGNTYMLLTNATLDREQWPKYTLTLLAQD
QGLQPLSAKKQLSIQISDINDNAPVFEKSRYEVSTRENNLPSLHLITIKAHADLGINQKVS
YRIQDSPVAHLVAIDSNTGEVTAQRSLNYEEMAGFEFQVIAEDSGQPMЛАSSVSVWVSSLDA
NDNAPEVVQPVLSDGKASLSVLVNASTGHLLVPIETPNGLGPAGTDPPPLATHSSRPFLTT
IVARDADSGANGEPLYSIRNGNEAHLFILNPTHGQLFVNVTNASSLIGSEWELEIVVEDQGS
PPLQTRALLRVMFVTSDHLDARKPGALMSMSMLTVICLAVLLGIFGLLALFMSICRTEK
KDNRAYNCREAESTYRQQPKRPQKHIQKADIHLVPVLRGQAGEPCEVGQSHKDVDKEAMMEA
GWDPCQLQAPFHLPQLYRTLNRNQGNQGAPAESREVLQDTVNLLFNHPRQRNASRENLNLP
QPATGQPRSPLKVAGSPTGRLAGDQGSEEAPQRPPASSATLRRQRHLNGKVSPEKESGPRQ
ILRSLVRLSVAAFAAERNPVEELTVDSPPVQQISQLLSSLHQGQFQPKPNHRGNKYLAKPGGS
RSAIPDTDGPSARAGGQTDPEQEEGPLDPEEDLSVKQLLEEELSSLLDPSTGLALDRLSAPD
PAWMARLSSLPLTTNYRDNVISPDAAATEEPRTFQTFGKAEAPELSPTGTRLASTFVSEMSSL
LEMLLERRSSMPVEAASEALRRLSVCGRTLSLDLATSAASGMKVQGDPGGKTGTEGKSRGSS
SSSRCL
```

Important features:

Signal peptide:

amino acids 1-13

Transmembrane domain:

amino acids 719-739

N-glycosylation site.

amino acids 415-418, 582-585, 659-662, 662-665 and 857-860

Cadherins extracellular repeated domain signature.

amino acids 123-133, 232-242, 340-350, 448-458 and 553-563

FIGURE 172

CGGACGGCGTGGCGGACGGTGGGGAGAGCCGAGTCCCGCTGCAGCACCTGGAGAAGG
CAGACCGTGTGAGGGGGCTGTGGCCCCAGCGTGTGGCCTGGGAGTGGAGTGGAG
GCAGGAGCCTCCTAACACTTCGCCATGAGTTCTCATCGACTCCAGCATCATGATTACCT
CCCAGATACTATTTTGATTGGCTTGTGGCTTCATGCGCCAATTGTTAAAGACTAT
GAGATACGTCAGTATGTTGTACAGGTGATCTCTCCGTGACGTTGCATTTCTGCACC
GTTGAGCTCATCTTGAATCTTAGGAGTATTGAATAGCAGCTCCGTTATTTCACT
GGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTCATGGTGCCTTTACATTGGC
TATTTATTGTGAGCAATATCCGACTACTGCATAAACAAACGACTGCTTTCTGTCTCTT
ATGGCTGACCTTATGTATTCTCTGGAAACTAGGAGATCCCTTCCCATTCTCAGCCAA
AACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGTGGTGTGATTGGAGTGACTCTC
ATGGCTTCTGGATTGGTGTCAACTGCCATACACTTACATGTCTTACTTCCT
CAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATA
TGATCATAAGCAAAAGAAAAGGATGGCAATGGCACGGAGAACATGTTCCAGAAGGGGAA
GTGCATAACAAACCATCAGGTTCTGGGAATGATAAAAGTGTACCACCTCAGCATCAGG
AAGTGAAAATCTTAACCTTACATTCAACAGGAAGTGGATGCTTGGAGAACATTAAGCAGGCAGC
TTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAACATGTTGGAAACACATT
AAGGGAAATATTTAATTCTGGTTACTTTCTCTATTTACTGTGTTGGAAAATTT
CATGGCTACCACATCAATTGTTGATCGAGTTGGAAACGGATCCTGTACAGAGGCA
TTGAGATCACTGTGAATTATCTGGGAATCCAATTGATGTGAAGTTGGTCCAACACATT
TCCTTCATTCTTGTGAAATAATCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTAC
CAAGTTCTTATGCCATCTCAGCAGTAAGTCCTCAATGTCATTGCTCTGCTATTAGCAC
AGATAATGGCATGTACTTGTCTCCTGTGCTGATCCGAATGAGTATGCCATTAGAA
TACCGCACCATAATCACTGAAGTCCTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGTT
TGATGTGATCTCCTGGTCAGCGCTCTCTAGCATACTCTCCTCTATTGGCTCACAAAC
AGGCACCAGAGAACGAAATGGCACCTTGAACTTAAGCCTACTACAGACTGTTAGAGGCCAGT
GGTTCAAAATTAGATATAAGAGGGGGAAAATGGAACCAGGGCCTGACATTATAAAC
AAACAAAATGCTATGGTAGCATTTCACCTCATAGCATACTCCTCCCGTCAGGTGATA
CTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACTAAGACAATCTCA
GCAGAGAGCATCCGTGGATATGAGGCTGGTAGAGGCCAGAGGCCAGAACACTAA
AGGTGAAAATACACTGGAACCTGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGT
AGGATTCCGTTTAAGGTTACATGGAAAAGGTTAGCTTGCCTGAGATTGACTCATT
AAAATCAGAGACTGTAACAAAAAAAAAAAAAGGGCGGCCGACTCTAGAGTCG
ACCTGCAGAAGCTTGGCCGACCTGGCCAACTGTTATTGCAGCTATAATG

FIGURE 173

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEI
LGVLNSSSSRYFHWKMNLCVILLILVFMVPFYIGYFIVSNIRLLHKQRLLFSCLLWLTMYFF
WKLGDPFPILSPKHGILSIEQLISRVGVIGVTLALLSGFGAVNCPTYMSYFLRNVTDTDI
LALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPSGFWMKSVTTSASGSENLTLIQ
QEVDALLELSRQLFLETADLYATKERIEYSKTFKGKYFNFLGYFFSIYCVWKIFMATINIVF
DRVGGKTDPVTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISS
SKSSNVIVLLAQLIMGMYFVSSVLLIRMSMPLYRTIITEVLGELQFNFYHRWFDVIFLVSA
LSSILFLYLAHKQAPEKQMAP

Important features:

Signal peptide:

amino acids 1-23

Potential transmembrane domains:

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398,
425-444

N-glycosylation sites.

amino acids 67-70, 180-183 and 243-246

Eukaryotic cobalamin-binding proteins

amino acids 151-160

FIGURE 174

CATGGGAAGTGGAGCCGGAGCCTTCCTTACACTGCCATGAGTTCTCATCGACTCCAGCA
TCATGATTACCTCCCNGANACTATTTTGATTTGGGTGGCTTCTCNGGCCAATGTT
TAAAGACTATGAGATACTGAGTATGTTGACNGGTGATCTTCTCCGTGACGTTGCCATT
CTTGCACCATGTTGAGCTCATCATCTTGAAATCTTNGGAGTATTGAATAGCAGCTCCGT
TATTTCACTGGAAAATGAACCTGTGTGAATTCTGCTGATCCTGGTTNTCATGGTGCCTTT
TTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCATAAACACGACTGCTTTTT
CCTGTCTCTTATGGCTGACCTTATGTATTCCAG

FIGURE 175

GTGTTGCCCTGGGGAGGGGAAGGGGAGCCNGGCCCTTCTAAATTTGGCCAAGGGTTCTTNTTGAATTCCGGGTTNNGNATACCTTCCCAGAAAATTTTTGGATTGGGGTAGNTT
TTTTTCATGCGCCAATTGTTAAAGACTATGAGATACGTCAGTATGTTGTACAGGTGATNTTNTCCGTGACGTTGCATTTCTTGCACCATGTTGAGCTCATCATNTTGAAATNTTAGGAG
TATTGAATAGCAGCTCCGTTATTTCACTGGAAAATGAAACCTGTGTGTAATTCTGCTGATCCTGGTTTCATGGTGCCTTTACATTGGCTATTTATTGTGAGCAATATCCGACTACTGCA
TAAACAAACGACTGCTTTTCTGTCTNTTATGGCTGACCTTATGTATTNTNTGGAAAANTAGGAGATCCCTTCCCATTCTC

卷之三

FIGURE 176

CTCGCCGAGGGATCGTCCCATGGCCGGGGCTCGAGCCGCACCCTTGGGGGCTCCGGGATTGCTACCTTT
TGGCTCCCTGCTCGAACTGCTCTTCACGGCTGCGCTTCAATCTGGACGTGATGGGGCCTGCGCAA
GGAGGGCGAGCCAGGCAGCCTCTCGGCTTCTCTGGCCCTGACCGGCAGTGAGCCCCGACCCCAGAGCTG
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGCAGCAGGCCAATCGCAGTGAGCCCCCTTCTGGCTTGG
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGACATCGACCAAGGGAGCTGATATGCAAAGGAAAGCAA
GGAGAACCACTGAGTGGTGGGAGTCAGTGTCGGAGCCAGGGGCTGGGGCAAGATTGTTACCTGTGACACCGATA
TGAGGAAGGCAGCGAGTGGACCAGATCTGGAGACGCCAGATATGATTGGTCGCTTGTGCTCAGCCAGGA
CCTGGCCATCGGGATGAGTTGGATGGGGATGGAAGTCTGTGAGGGACGCCAAGGCCATGAAACAAATT
TGGGTTCTGCCAGCAGGGCACAGCTGCCCTTCTCCCCTGATAGCCACTACCTCCCTTTGGGGCCCAAGGAAC
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGACAGGGCTCAGGGACCTGGCACACCTGGACACGG
TCCCTACGAGGCGGGGGAGAGAACGGAGCAGGACCCCGCCTCATCCGGTCCCTGCCAACAGCTACTTGGCTT
CTCTATGACTCGGGAAAGGTCTGGTGCAGAAGAGCTGAGCTTGTGGCTGGAGCCCCCGCGCCAAACCA
CAAGGGTGTGTCATCTGCGCAAGGACAGGCCAGTCCCTGGTGCCTGGGGTATGCTGCTGGGGAGCG
CCTGACCTCCGGCTTGGCTACTCACTGGCTGTGGCTGACCTAACAGTGATGGCTGGCCAGACCTGATAGTGG
TGCCCCCTACTTCTTGAGCGCCAAGAAGAGCTGGGGGTGCTGTATGTGACTTGAACCAGGGGTCACTG
GGCTGGGATCTCCCTCTCCGGCTGCGGCTCCCTGACTCCATGTTGGGATCAGCCTGGCTGTGCTGGGG
CCTCAACCAAGATGGCTTCCAGATATTGAGCTGGGATGGGAAAGTCTTACATCTACCA
TGGGAGCAGCCTGGGGTTGCGCCAAACCTCAAGGTGCTGGAGGGCGAGGCTGGGATCAAGAGCTTCGG
CTACTCCCTGTCAGGCAGCTGGATGGGAAACCAATACCCCTGACCTGCTGGGCTCCCTGGCTGACAC
CGCAGTGCTCTCAGGGCCAGACCCATCTCCATGCTCCCATGAGGTCTATTGCTCCACGAAGCATGACCT
GGAGCAGCCAACTGTGCTGGGGCACTGGCTGTGGACCTAAGGGCTGTTCAGGCTACATTGAGCTCC
CAGCAGCTATAGCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCCGGAGGCTCCGGGGCAGGT
TCCCCGTGACGTTCTGAGCCTAACCTGGAAAGAACCCAGCACCAGGCCTGGGACCCGGTGTGGCTGAAGCA
CCAGCATGACCGAGTCTGTGAGACGCCATGTTCCAGCTCAGGAAATGTCAAAAGACAAGCTCGGGCATTGT
AGTGACCTGTGCTACAGTCTCCAGACCCCTCGGCTCCGGCAACGGCTCTGGCCAGGGCTGCTCCAGTGG
CCCCATCTCAATGCCAACAGCCAGCAGCAGGGCAGAGATCCTCTGAAGCAAGGCTGTGGTAAGA
CAAGATCTGCCAGAGCAATCTGAGCTGGTCCACGCCCCTCTGTACCCGGGTAGCGACACGGAAATTCCAACC
TCTGCCATGGATGTGGATGGAACAACAGCCCTGTTGACTGAGTGGGAGCCAGTATTGGCTGGAGCTGAT
GGTACCAACCTGCCATGGACCCAGCCAGCCCCAGGCTGATGGGATGATGCCATGAAGCCAGCTCTGGT
CATGCTCCTGACTCACTGCACTACTCAGGGTCCGGGCCCTGGACCTGCGGAGAAGCCACTTGCCCTGCTCAA
TGAGAATGCCCTCCATGTTGAGTGTGAGCTGGGGAAACCCATGAAGAGAGGTGCCAGGTACCTTCTACCTCAT
CCTTAGCACCTCCGGATCAGCATTGAGACCAGGAACCTGGAGGTAGAGCTGCTGTGGCCAGATCAGTGAGCA
GGAGCTGCACTCAGTCTGACGAGCCGTGCTCATTGAGCTGCCACTGTCATTGAGGAATGGCCATTCC
CCAGCAACTCTTCTCTGGTGTGGAGGGCGAGAGAGCCATGCAGTCTGGGATGTGGCAGCAAGGT
CAAGATGAGGTCACGGTTCCAACCAAGGCCAGTCGCTCAGAACCCCTGGCTCTGCCCTCTCAACATCATGTG
GCCCTAGAGATTGCCAATGGGAAGTGTTGCTGACCCATGCAAGGTTGAGCTGGAGGGCAGGGCAGGGGCTGG
GCAGAAAGGGCTTGTCTCCAGGCCAACATCTCCACCTGGATGTGACAGTAGGGATAGGAGGCGGG
GCTGGAGGCCACTGAGCAGCAGGCCAGGGCAGAGCTGGTGTGGAGGCCAGCATGTCCTGGTGGCCAGTGTCTCTGC
TGAGAAGAAGAAAACATCACCCTGGACTGCGCCGGGGCACGGCAACTGTGTTGAGCTGGCCACTCTA
CAGCTTGTACCGCGCGGCTGTGCTGACATGTCATGTCGGGGCAACATCAGTGAGCTGGAGGAGTACTCAGC
TGTGAAGTCCCTGGAAGTGATGTTGCTGGGCAACATCAGTGAGCTGGTAAAGAACATTGATGCTCCGAGA
TGCCTCCACAGTGATCCCAGTGATGGTAACTCTGGACCCCATGGTGTGGCAGAAGGAGTGCCTGGTGGGT
CATCCTCCTGGCTGACTGGCTGGCTGCTGGTGTGACTGCTGGTGTGCTGGACAGATGGCTCTGGAAAGATGGGATTCTT
CAAACGGGCGAAGCACCCCGAGGCCACCGTGGCCACTGAGATGGCTCTGGGAGATCTCCGGAAAGACCGACAGCA
GTTCAAGGAGGAGAACAGGGCACCATCTGAGGAACAACTGGGGCAGCCCCGGGGAGGGCCAGGGCAG
CCCCATCTGGCTGCTGACGGGCATCCCGAGCTGGGCCAGGATGGGCATCAGGGCAGGGCAGGGCTGGG
CATGTCCCAGCCTGGCTGTGGCTGCCCTCCATCCCTTCCAGAGATGGCTCTGGGATGAAGAGGGTAGAGT
GGGCTGCTGGTGTGCATAAGATTTGGCAGGATCGGCTTCTCAGGGGACAGACCTCTCCACCCAAAGAAC
TCCTCCCACCCAAACTTCCCTTAGAGTGCTGTGAGATGAGCTGGGAAATCAGGGACAGGGCAGGGCAGGG
TGAGAAGGGCAGGGTGTCTGATGCAAAGGTGGGAGAACGGATCCTAATCCCTTCTCTCCATTCACTG
GTAACAGGACCCAAAGGACCTGCTCCCCGGAAAGTGCCTTAACCTAGAGGGTGGGGAGGGTAGTGTCACTGA
CTCAGGCTGCTCTCTAGTTCCCTCATCTGACCTTAGTTGCTGCCATCAGTCAGTGTTCTGCTGG
TTCGCTATTAAAAAATATTGAGAACAAAAAAAAAAAAAA

FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL
QPRPQSLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL
GVSVRSGPGGKIVTCARYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDGGEWFCEG
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGTARVELCAQGSADLAHLDGPYEA
GGEKEQDPRLIPVPANSYFGFSIDSGKGLVRAEELSFVAGAPRANHKGAVVILRKDSASRLV
PEVMLSGERLTSGFGYSLAVADLNSDGWPDLIVGAPYFFERQEEELGGAVYVYLNQGGHWAGI
SPLRLCGSPDSMFGISLAVLGDLNQDGFPDIAVGAPFDGDGKVFYHGSSLGVVAKPSQVLE
GEAVGIKSFYSLSGSILDMDGNQYPDLLVGSЛАDTAVLFRARPILHSHEVSIAPRSIDLEQ
PNCAGGHSVCVDLRVCFSYIAVPPSYSPVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK
HQASGTVWLKHQHDRVCVDAMFQLQENVKDKLRAIVTLSYLQTPRLRRQAPGQGLPPVAP
ILNAHQPSTQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRSDTEFQQLPMVDGTTALFA
LSGQPVIGLELMVTNLPSPDAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSN
ENASHVECELGNPMKRGAQVTFYLILSTSGISIETTELEVELLLATISEQELHPVSARARVF
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVYEVTVSNQGQSLRTLGSFLNIM
WPHEIANGKWLLYPMQVELEGGQGPGQKGLCSPRPNILHLDVDSRDRRRRELEPPQQEPGE
RQEPMSSWWPVSSAEKKNITLDCARGTANCVVFCPLYSFDRAAVLHWGRLWNSTFLEEY
SAVKSLEVIVRANITVKSSIKNIMLRDASTVIPVMVYLDPMMAVVAEGVPWWVILLAVLAGLL
VIALLVLLWKMGGFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGSPRREGP
DAHPILAADGHPELGPDPGHPGPFTA

Important features:

Signal peptide:

amino acids 1-33

Transmembrane domain:

amino acids 1040-1062

N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

2023-09-27 10:07:00

FIGURE 178

CGCGCCGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGCGTGCAGCAGCTCCAGA
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTCCAGCAGGAGCTGGAGCACAGTGCTGGCT
CACAAACAAGTGCTCAAGGTGTCAAGCGTACTGTGTGTGCAGCCGCTTGGTGCAGTCA
GTCTCTCGCAGCTGCCCGCGTGGCTGCAGCCGGGGCGGTGGACGGCGTAATTTCT
TGGATGATAAACAAATGGCTACCCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC
AAATTCCGAGACGAAGTAGAGGATGATTATTCGCACTTGGAGTCCAGGAAAACCCCTCGA
TCAGGCTTAGATCCAGCTAAGGATCCATGCTAAAGATGAAATGTAGTCGCCATAAGTAT
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG
AAAGAACAGGAGTAGACCATAGGCAGTGGAGGGTCCCATTATCCACCTGCAAGCAGTG
CCCAGTGGCTATCCAGCCCTGTTGGTTCAGATGGTCATAACCTACTCTTTCAGTGCA
AACTAGAATATCAGGCATGTGCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC
CCATGTCCTTCAGATAAGCCCACCAAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT
GGAGTTCAGGAAGTGGCAAACAGATTGGGACTGGTCAAGGCCCTCATGAAAGTGGAA
GTCAAAACAAGAACAAAAACATTGCTGAGGCCTGAGAGAACAGATTGATAACCAGCATC
TTGCCAATTGCAAGGACTCAGTGGCTGGATGTTAACAGACTTGATAACAAACTATGACCT
GCTATTGGACCAGTCAGAGCTCAGAACGATTACCTTGATAAGAACAGTGTACCAAGG
CATTCTCAATTCTTGACACATACAAGGACAGTTAATATCTAATAATGAGTGGTGTAC
TGCTTCCAGAGACAGCAAGACCCACCTGCCAGACTGAGCTCAGCAATATTCAAGCGGCA
AGGGTAAAGAACAGCTCAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC
CAACACAATGTCATGGCAGTGGACAGTGCTGGTGTGACAGATATGAAATGAAGTC
ATGGGATCCAGAATAATGGTGTGCAAGATTGTGCTATAGATTGAGATCTCCGGAGATTT
TGCTAGTGGCGATTTCATGAATGGACTGATGATGAGGATGATGAAGAGACGATATTGAAATG
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGATGATGATGATGGTGGTGTGAC
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAATTCTACATTTCTAATATTAA
CAAAAATGATAGCCTATTTAAATTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA
CATATATTGTATAATTATTTGAAAAATTGCAGCTAAAGTTAGAACTTTATGTTAAAT
AAGAACATTTGCTTGAGTTTATATTCCCTACACAAAAGAAAATACATATGCAGTCTA
GTCAGACAAAATAAGTTGAAGTGCTACTATAATAATTTCACGAGAACAAACTTGT
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTGGATCGTACATGTTAATTGGTAAAG
ATAATTCTAAGTGAATTAAAATAAAATTGACCTGGTCTTAAGGATTAGG
AAAAATATGCATGCTTAATTGCATTCCAAAGTAGCATTTGCTAGACCTAGATGAGTCAG
GATAACAGAGAGATAACCACATGACTCCAAAAAAAAAAAAAA

FIGURE 179

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49829
><subunit 1 of 1, 436 aa, 1 stop
><MW: 49429, pI: 4.80, NX(S/T): 0
MLKVS A V L C V C A A A W C S Q S L A A A A V A A A G G R S D G G N F L D D K Q W L T T I S Q Y D K E V G Q W N K F R
D E V E D D Y F R T W S P G K P F D Q A L D P A K D P C L K M K C S R H K V C I A Q D S Q T A V C I S H R R L T H R M K E A
G V D H R Q W R G P I L S T C K Q C P V V Y P S P V C G S D G H T Y S F Q C K L E Y Q A C V L G K Q I S V K C E G H C P C P
S D K P T S T S R N V K R A C S D L E F R E V A N R L R D W F K A L H E S G S Q N K K T K T L L R P E R S R F D T S I L P I
C K D S I L G W M F N R L D T N Y D L L L D Q S E L R S I Y L D K N E Q C T K A F F N S C D T Y K D S L I S N N E W C Y C F Q
R Q Q D P P C Q T E L S N I Q K R Q G V K K L L G Q Y I P L C D E D G Y Y K P T Q C H G S V G Q C W C V D R Y G N E V M G S
R I N G V A D C A I D F E I S G D F A S G D F H E W T D D E D D E D D I M N D E D E I E D D D E D E G D D D G G D D H D V Y I
```

Important features:

Signal peptide:

amino acids 1-16

Leucine zipper pattern.

amino acids 246-267

N-myristoylation sites.

amino acids 357-362, 371-376 and 376-381

Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352

FIGURE 180

FIGURE 181

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52196
><subunit 1 of 1, 229 aa, 1 stop
><MW: 26017, pI: 4.73, NX(S/T): 0
MGDKIWLPPVLLAALPPVLLPGAAAGFTPSLSDSDFTFTLPAGQKECFYQPMPLKASLEIEY
QVLDGAGLDIDFHLASPEGKTLVFEQRKSDGVHTVETEVGDYMFCDNTFSTISEKVIFFEL
ILDNMGEQAQEQQEDWKKYITGTDILDLMKLEDILESINSIKSRLSKSGHIQILLRAFEARDRN
IQESNFDRVNFWSMVNLVVMVVSAIQVYMLKSLFEDKRKSRT
```

Important features:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 195-217

N-myristoylation site.

amino acids 43-48

Tyrosine kinase phosphorylation site.

amino acids 55-62

FIGURE 182

CCATCCCTGAGATCTTTATAAAAAACCCAGTCTTGCTGACCAGACAAAGCATAACCAGAT
CTCACCAAGAGAGTCGAGACACTATGCTGCCTCCATGCCCTGCCAGTGTGTCCTGGATG
CTGCTTCCTGCCTCATTCTCCTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAACGTGC
CTCTCCACGGATCAGCTGTCAAAGGCTCAAGGCCTATGGCTCCCCCTGCTATGCCTTGT
TTTGTCACCAAAATCCTGGATGGATGCAGATCTGGCTGCCAGAAGCGGCCCTGGAAAAA
CTGGTGTCTGTGCTCAGTGGGCTGAGGGATCCTCGTGTCCCTGGTGAGGAGCATTAG
TAACAGCTACTCATAACATCTGGATTGGCTCCATGACCCCACACAGGGCTCTGAGCCTGATG
GAGATGGATGGAGTGGAGTAGCACTGATGTGATGAATTACTTGCATGGAGAAAAATCCC
TCCACCATCTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTG
GAAAGATTATAACTGTGATGCAAAGTTACCCATGTCTGCAAGTTCAAGGACTAGGGCAGGT
GGGAAGTCAGCAGCCTCAGCTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGAC
TCACCCCTGGAAGAGAATATTCTCCCCAACTGCCCTACCTGACTACCTTGTATGATCCTCC
TTCTTTTCCTTTCTCACCTCATTTCAGGCTTCTGTCTCCATGTCTTGAGATC
TCAGAGAATAATAATAAAAAATGTTACTTATAAAAAAAAAAAAAAAA

FIGURE 183

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965
<subunit 1 of 1, 175 aa, 1 stop
<MW: 19330, pI: 7.25, NX(S/T): 1
MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWM
DADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSS
TDVMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWDYNCDAKLPYVCKFKD
```

Important features:

Signal peptide:

amino acids 1-26

C-type lectin domain signature.

amino acids 146-171

FIGURE 184

CCAGTCTGTCGCCACCTCACTTGGTGTCTGCTGCCCCGCCAGGCAAGCCTGGGTGAGAGC
ACAGAGGAGTGGCCGGGACCATGGGGGGACGCCGGCTGGCGCTCCTGGCGCTGGTGC
TGCCTGCGGAGAGCTGGCGCCGGCCTGCGCTGCTACGTCTGTCCGGAGCCCACAGGAGTGT
CGGACTGTGTCACCATGCCACCTGCACCAACGAAACCATGTGCAAGACCACACTCTAC
TCCCAGGAGATAGTGTACCCCTCCAGGGGACTCCACGGTGACCAAGTCCTGTGCCAGCAA
GTGTAAGCCCTCGGATGTGGATGGCATGGCCAGACCCCTGCCGTGTCCTGCTGCAATACTG
AGCTGTGCAATGTAGACGGGCGCCGCTCTGAACAGCCTCCACTGCCGGCCCTCACGCTC
CTCCCACTCTTGAGCCTCCGACTGTAAGAGTCCCCGCCACCCCCATGCCCTATGCCGGCCCA
GCCCGAATGCCCTGAAGAAGTGCCCTGCAACCAGGAAAAAAAAAAAAAA

FIGURE 185

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56405
<subunit 1 of 1, 125 aa, 1 stop
<MW: 13115, pI: 5.90, NX(S/T): 1
MRGTRLALLALVLAACGELAPALRCYVCPEPTGVSDCVTIATCTTNETMCKTLYSREIVYP
FQGDSTVTKSCASKCKPSDVGIGQTLPVSCCNTELCNVDGAPALNSLHCGALTLLPLLSLRL
```

Important features:

Signal peptide:

amino acids 1-17

N-glycosylation site.

amino acids 46-49

FIGURE 186

CTGCAGTCAGGACTCTGGACCGCAGGGGCTCCGGACCTGACTCTGCAGCCGAACCGGC
ACGGTTCTGGGGACCCAGGCTTGC~~AA~~AGTGACGGTCATTTCTCTTCTCCCTCTT
GAGTCCTTCTGAGATGATGGCTCTGGGCCAGCAGGGAGCTACCCGGTCTTGTGCGATGG
TAGCGGCCGCTCTCGCCGCCACCCCTCTGCTGGAGTGAGCGCCACCTGAACTCGGTTCTC
AATTCCAACGCTATCAAGAACCTGCCACCGCTGGCGCTGCGGGCACCCAGGCTC
TGCAGTCAGGCCGCCGGAAATCCTGTACCCGGCGGAATAAGTACCAAGACCATTGACA
ACTACCAGCCGTACCCGTGCGCAGAGGACGAGGAGTGCGGCACGTGAGTACTGCGCTAGT
CCCACCCCGCGAGGGACGCAGCGTCAAATCTGTCTCGCCTGCAGGAAGCGCGAAAACG
CTGCATGCGTCACGCTATGTGCTGCCCGGAATTACTGCAAAATGGAATATGTGTGCTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTGGATGGTATTCCAGAAGAACCAACCTTGTCTCAAAATGTATCACACCAA
AGGACAAGAAGGTTCTGTTGCTCCGGTCATCAGACTGTGCCTCAGGATTGTGTTGCTA
GACACTTCTGGTCCAAGATCTGTAACACCTGCTGAAAGAAGGTCAAGTGTACCAAGCAT
AGGAGAAAAGGCTCTCATGGACTAGAAATATTCCAGCGTTACTGTGGAGAAGGTCTGTC
TTGCCGGATACAGAAAGATCACCATCAAGCCAGTAATTCTCTAGGCTTCACACTGTCAGA
GACACTAAACCAGCTATCAAATGCAGTGAACTCCTTATATAATAGATGCTATGAAAACC
TTTATGACCTTCATCAACTCAATCCTAAGGATATAACAGTTCTGTTAGGTTAGTTAAC
TCCAATAACACCTTCAAAACCTGGAGTGTAAAGAGCTTGTCTTATGGAACCTCCCTG
TGATTGCAGTAAATTACTGTATTGAAATTCTCAGTGTGGCACTTACCTGAAATGCAATGA
AACTTTAATTATTTCTAAAGGTGCTGCACTGCCTATTCTCTTGTATGAAATT
TTGTACACATTGATTGTTATCTGACTGACAAATATTCTATATTGAACGTAAATCATT
TCAGCTTATAGTTCTAAAAGCATAACCCCTTACCCATTAAATTCTAGAGTCTAGAACGCA
AGGATCTCTGGAATGACAAATGATAGGTACCTAAATGTAACATGAAAATACTAGCTTATT
TTCTGAAATGTAATCTTAAATTATATTCCCTTAGGCTGTGATAGTTTGAA
AATAAAATTAAACATTAAAAAAAAAAAAAA

FIGURE 187

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57530
<subunit 1 of 1, 266 aa, 1 stop
<MW: 28672, pI: 8.85, NX(S/T): 1
MMALGAAGATRVFVAMVAAALGGHPLLGVSATLNSVLNSNAIKNLPPPLGGAAGHPGSAVSA
APGILYPGGNKYQTIDNYQPYP PCAEDEECGTDEYCASPTRGGDAGVQICLACRKRRKRCMRH
AMCCPGNYCKNGICVSSDQNHF RGEIEETITESFGNDHSTLDGYSRRTTLSSKMYHTKGQEG
SVCLRSSDCASGLCCARHFWSKICKPVLKEGQVCTKHRRKGSHGLEIFQRCYC GEGLSCRIQ
KDHHQASNSSRLHTCQRH
```

Important features:

Signal peptide:

amino acids 1-23

N-glycosylation site.

amino acids 256-259

Fungal Zn(2)-Cys(6) binuclear cluster domain

amino acids 110-126

FIGURE 188

TGTGTTCCCTGCAGTCAGAATTGGGACNGCAGGGTTCCGGACCTGATTTCAGCGGA
ACGGGAAGGTTTGTGGGACCCAGGTTGAAATGACGGTCATTTTTCTCCTTCNG
GAGTCCTNTGAGANGATGGTTTGGGCCAGCGGGAGCTAACCGGTTTGTNGCGATG
GTAGCGCGGTTTGGCGGCCACCTNTGCTGGAGTGAGCGCCACCTGAATCGGTTTC
AATTCCAACGNTATCAAGAACCTGCCACCGNTGGCGGCCGCTGGGGCACCCAGGNTT
TGCAGTCAGGCCGCGCCGGAAATCCTGTACCCGGCGGGAAATAAGTACCAAGACCATTGACA
ATTACCAGCCGTACCGTGCGCAGAGGACGAGGAGTGCGGCACTGATGAGTACTGCGCTAGT
CCCACCCGCGGAGGGGANGCGGGCGTGCAGGAAATNTGNTNGCCTGCAGGAAGCGCCGAAACG
CTGCATGCGTCANGCTATGTGCTGCCCGGGAAATTACTGCAAAATGGAATATGTGTNTT
CTGATCAAAATCATTCCGAGGAGAAATTGAGGAAACCATCACTGAAAGCTTGGTAATGAT
CATAGCACCTTGGATGGG

FIGURE 189

GAGGAACCTACCGGTACCGGCCGCGCCTGGTAGTCGCCGTGCGCTGCACCTCACCAATCCCGTGCGCCGG
CTGGGCCGTCGGAGAGTGCCTGCTTCTCTGCACGCCGTGCTGGCTGCCAGGCCGGTCCGCC
GGGTTGAGGATGGGGAGTAGCTACAGGAAGCAGCCCGCATGGCAAGGTATTTTGAGGAAATGAAAAGGA
AGTATTAGAAATGAGCTGAAGACATTACAGATAATATTGGGACAGATTGTGATGCTGATTACCC
TGAAGTAATGTAGACAGAAGTCTCAAATTGATATTACATCAACTGGAACCAAGCAGTGAATCTTAATGTCAC
TTAAATCAGAACATTGATAAGAAGAGAATGGGAGTCTGGTTAAATAAAGATGACTATATCAGAGACTTGAAGA
GATCATTCTCTGTTCTGATAGTGTATGGCATTTAGTGGCACAGATCAGGATTTACAGTTACTTGG
AGTGTCCAAACTGCAAGCAGTAGAGAAATAAGACAAGCTTCAAGAAATTGGCATTGAAGTTACATCCTGATAA
AAACCGAATAACCCAAATGCCACATGGCATTTTAAAAATAATAGAGCATATGAAGTACTCAAAGATGAAGA
TCTACGGAAAAGTATGACAAATATGGAGAAAAGGGACTTGAGGATAATCAAGGTGCCAGTATGAAAGCTGGAA
CTATTATCGTTATGATTTGGTATTTATGATGATGATCCTGAAATCATAACATTGAAAGAAGAGAATTGATGC
TGCTGTTAATTCTGGAGAACTGTGGTTGTAAATTAACTCCCCAGGCTTCACTGCCATGATTTAGCTCC
CACATGGAGAGACTTGTCAAAGAAGTGGATGGGTTACTTCAATTGGAGCTTTAATGTGGTGTGATGATAGAAT
GCTTGCAGAATGAAAGGAGTCAACAGCTATCCCAGTCTCATTTTCCGGTCTGGAATGGCCCACTGAAATA
TCATGGAGACAGATCAAAGGAGAGTTAGTGAATTGCAATGCAGCATGTTAGAAGTACAGTGCACAGAACATTG
GACAGGAAATTGTCACCTACAAACTGCTTGTCTGGTATTGGCTGGTGTACACTTTGTTCAA
AGGAGGAGATTGTTGACTTCACAGACACGACTCAGGCTTACTGGCATGTTGTTCTCAACTCATTGGATGCTAA
AGAAATATAATTGGAAGTAATAACATAATCTCCAGATTGAACTACTTCGGCAAACACACTAGAGGATCGTT
GGCTCATCAGGGCTGTTATTTCATTTGGAAAAAAATGAAATTCAAATGATCCTGAGCTGAAAAAAACT
AAAAACTCTACTAAATGATCATATTCAAGTTGGCAGGTTGACTGTTCTGCAACAGACATCTGAGTAA
TCTGTATGTTTCAGCCGCTCTAGCAGTATTAAAGGACAAGGAAACAAAGAATATGAAATTCAATGGAA
GAAGATTCTATATGATATACTTGCCTTGGCAAAGGAGTGTGAATTCTCATGTTACACCGCTGGACCTCAA
TTTCTGCCAATGACAAAGAACATGGCTTGTGATTCTTCTGGCTGGCTCACCAGTCGAGCTTACT
ACCAAGAGTTACGAAGAGCATCAAATCTTATGGTCAAGGTTGACTGTTCAACTAGATTGACAGTTCATGA
GGGACTCTGTAACATGTATAACATTCAAGGCTTATCCAACACAGTGGTATTCAACAGCTCAACATTGAGTA
TGAAGGACATCACTGCTGAACAAATTGGAGTTCATAGAGGATCTATGAAATTCTTCAGTGGCTCCCTTAC
ACCCACCACTTCAACGAACTAGTACACAAAGAAAACACAACGAAGTCTGGATGGTGTGATTCTATTCTCCGTG
GTGTCATCCTGGCAAGTCTTAATGCCAGAATGGAAAAGAATGGCCGGACATTAACAGTGGACTGATCAACGTGG
CACTAGATTGCCAACAGTATCATTCTTGTGCCAGGAAACGTTCAAAGATAACCTGAGATAAGATT
TCCCCAAAATCAAATAAGCTTACAGTATCACAGTTACAATGGTGAATAGGGATGTTATTCCCTGAGAAT
CTGGGTCTAGGATTTCACCTCAAGTATCCACAGATCTAACACCTCAGACTTCACTGAAAAAGTTCTACAAGG
AAAAATCATTGGGTGATTGATTCTATGCTCTGGTGTGGACCTGGCAGAATTGCTCCAGAATTGAGCT
CTTGGCTAGGATGATTAAGGAAAAGTGAAGCTGGAAAAGTAGACTGTCAGGCTTATGTCAGACATGCCAGAA
AGCTGGGATCAGGGCTATCCAACAGTTAAGTTATTCTACAGAAAGAGAAATTCAAGAAGAGCA
GATAAATACCAAGAGATGCAAAAGCAATCGCTGCCAATAAGTAAAAAGGAAACTCTCCGAAATCAAGGCAA
GAGGAATAAGGATGAACTTGTATAATTGAGGATGAAAGAAAAGTTAAAGAAATTCTGACAGATGACATCAG
AAGACACCTATTAGAATGTTACATTATGATGGGAATGAATGAACATTATCTTAGACTTGCAGTTGACTGCCA
GAATTATCTACAGCACTGGTAAAGGAGGCTGCAAACATTCTGTAAGGCGGTTATAAAATATT
GACTTGCAGGCTATAATATGTTCAATGAGAACAAGAATAGAGTCATCAAGTATTCTTGTATTGCT
TTAACACCTTAAAAAATATTAAACGATTCTAGCTCAGAGCCATACAAAGTAGGCTGGATTCACTG
GACCATGATTGCTGCTCCCTCGACGGACTTATAATGTTCAGGGCTGGCTGAAACATGAGCTGCTGTGCT
ATCTACATAATGCTAAGTTGTATAAGTCCACTTCCCTTCAGGTTGGCTGACCTGAAAGAGGTAAC
TAGTTGGTCACTGTTCTCTAAATGCTATCCCTAACCATATATTATATTGTTTAAACACCCAT
GATGTGGCACAGTAAACAAACCTGTTATGCTGATTATTATGAGGAGATTCTCATTGTTCTTCTCTCA
AAGGTTGAAAAAAATGCTTTAATTGGTACAGCCGAGAAACAGTGCAGCAGTATATGTCACACAGTAAGTACAC
AAATTGAGCAACAGTAAAGTGCACAAATTCTGAGTTGCTGATCATCCAGGAAACCTGAGGGAAAAAATTA
TAGCAATTAACTGGCATTGAGGATCTAAATATGTTACAGTATTAGAGTTCTATATTAAAGATATA
TGTGTTCATGTTATTCTGAAATTGCTTCAAGAAATTCCCACAGTATGTTGAGTTGAGGCACTAATAT
TTACATATTGCTTCTGAACTTGTGTTGACCTGTTACCTTATTACATTGGGTTTTCTTCACTAGTTGG
TTTTCACTCCTGTCAGTCTATTATTAACTGAGGAAATTACTTACAGGGTGTGTTACTGAGCTT
AATGATACTGTTAGGTTATTCCAGTTACTGTTACAGTGCAGGGCTGCCATTGAGATAAAATTGACATAATA
ACTGAGTTATTGTTATAAGGAAATCAAGTATATAAAATCTAGGAAAGGATCTTCACTGTTGTTAGA
CTCAAGAATCACAATTGTCAGTAAACATGTTAGTTAGTTATAATTCAAGAGTGTACAGAATGGTAAAATT
CCAATCAGTCAAAGAGGCTAATGAATTAAAGGCTTGCACACTTTCAAAAAAAAAAAAAAA

FIGURE 190

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56439
<subunit 1 of 1, 747 aa, 1 stop
<MW: 86127, pI: 7.46, NX(S/T): 2
MGVWLNKDDYIRDLKRIILCFLIVYMAILVGTQDFYSLLGVSKTASSREIRQAFKKLALKL
HPDKNPNNPNAHDFLKINRAYEVLKDEDLRKKYDKYGEKGLEDNQGGQYESWNYYRYDFGI
YDDDPPEIITLERREFDAAVNSGELWFVNFSYSPGCSHCHDLAPTWRDFAKEVDGLLRIGAVNC
GDDRMLCRMKGVNNSYPSLFIGRSGMAPVKYHGDRSKESLVSFAMQHVRSTVTELWTGNFVNS
IQTAFAAGIGWLITFCSKGGDCLTSQTRLRLSGMLFLNSLDAKEIYLEVIHNLPDFELLSAN
TLEDRLAHRWLLFFHFGKNENSNDPELKKLKTLKNDHIQVGRFDCCSSAPDICSNLYVFQP
SLAVFKGQGTKEYEIHGKKILYDILAFAKESVNHSVTLGPQNFPANDKEPWLVDFFAPWC
PPCRALLPELRRASNLLYGQLKFGTLDCTVHEGLCNMYNIQAYPTTVVFNQSNIHEYEGHHS
AEQILEFIEDLMNPSVVSLLPTTFNELVTQRKHNEVWMVDYSPWCHPCQVLMPEWKRMA
LTGLINVGSIDCQQYHSFCAQENVQRYPEIRFFPPKSNKAYQYHSYNGWNRDAYSRLRIWGLG
FLPQVSTDLPQTFSEKVLQGKNHWVIDFYAPWCGPCQNFAPFELLARMIKGKVKAGKVDC
QAYAQTCAKAGIRAYPTVKFYFYERAKRNFQEEQINTRDAKAI
AALISEKLETLRNQGKRNKDEL

Important features:

Endoplasmic reticulum targeting sequence.

amino acids 744-747

Cytochrome c family heme-binding site signature.

amino acids 158-163

Nt-dnaJ domain signature.

amino acids 77-96

N-glycosylation site.

amino acids 484-487

FIGURE 191

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAA
GCCATGAACATCATCCTAGAAATCCTCTGCTCTGATCACCATCATCTACTCCTACTTGGAGTCGGTGGTGAAGTTTCATTCTCAGAGGAGAAAATCTGTGGCTGGGAGATTGTTCTCATTA
TTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTTATGAATTGCAAAACGACAGAGC
ATATTGGTTCTGTGGATATTAATAAGCGCGGTGGAGGAAACTGCAGCTGAGTGCCGAAA
ACTAGGCGTCACTGCGCATCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCT
CTCTAAATCAGGTGAAGAAAGAAGTGGGTATGTAACAATCGTGGTGAATAATGCTGGGACA
GTATATCCAGCCGATCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAA
CATCCTAGGACATTTGGATCACAAAGCACTTCTCATCGATGATGGAGAGAAATCATG
GCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGATTCCCTACCTCATCCCCATAT
TGTCCAGCAAATTGCCGCTGTTGGCTTCACAGAGGTCTGACATCAGAACTTCAGGCCTT
GGGAAAAACTGGTATCAAAACCTCATGTCCTGCCAGTTTGTGAATACTGGGTTACCA
AAAATCCAAGCACAAGATTATGCCCTGTATTGGAGACAGATGAAGTCGAAGAAGTCTGATA
GATGGAATACTTACCAATAAGAAAATGATTTGTTCCATCGTATATCAATATCTTCTGAG
ACTACAGAAGTTCTCCTGAACGCCCTCAGCGATTTAAATCGTATGCAAGAATATTCAAT
TTGAAGCAGTGGTGGCCACAAATCAAATGAATAAAAGCTCCAGCCAGAGATG
TATGCATGATAATGATATGAATAGTTGCAATCAATGCTGCAAAGCTTATTCACATTTT
TCAGTCCTGATAATATTAACATTGGTTGGCACTAGCAGCAGTCAAACGAACAAGATTA
ATTACCTGTCTCCTGTTCTCAAGAATATTCACGTAGTTTCAAGGTCTGTTTCCCT
TCATGCCCTTAAACCTCTGTGCTTACATAAACATACTTAAAGGTTTCTTAAGATAT
TTTATTTTCCATTAAAGGTGGACAAAGCTACCTCCCTAAAGTAAATACAAAGAGAACT
TATTTACACAGGGAAAGGTTAAGACTGTCAGTAGCATTCCAATCTGTAGCCATGCCACAG
AAATCAACAAGAACACAGAACATGAGTGACAGCTAACAGAGATCAAGTTCAGCAGGCAGCTT
ATCTCAACCTGGACATATTTAAAGATTGAGCATTGAAAGATTTCCCTAGCCTCTTCTGTAT
TCATTAGCCAAAACGGTGCAACTCTATTCTGGACTTATTACTGATTCTGTCTTCTGTAT
AACTCTGAAGTCCACCAAAAGTGGACCCTCTATATTCCTCCCTTTATAGTCTTATAAGA
TACATTATGAAAGGTGACCGACTCTATTAAATCTCAGAATTAAAGTTCTAGCCCCATGA
TAACCTTTCTTGTAATTATGCTTCATATATCCTGGTCCAGAGATGTTAGACAAT
TTTAGGCTAAAAATTAAAGCTAACACAGGAAAGGAACGTACTGGCTATTACATAAGAAA
CAATGGACCAAGAGAAGAA

400-2300-3000-4000

FIGURE 192

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409
<subunit 1 of 1, 300 aa, 1 stop
<MW: 33655, pI: 9.31, NX(S/T): 1
MNIILEILLLITIIYSYLESLVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFARKQSI
LVLWDINKRGVEETAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKKEVGDVTIVVNNAGTV
YPADILLSTKDEEITKTFEVNILGHFWITKALLPSMMERNHGHIVTVASVCGHEGIPYLIPYC
SSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFVNTGFTKNPSTRLWPVLETDEVVRSЛИD
GILTNKKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK
```

Important features:

Signal peptide:

amino acids 1-19

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 30-33 and 58-61

Short-chain alcohol dehydrogenase family protein

amino acids 165-202, 37-49, 112-122 and 210-219

FIGURE 193

CGGCGGGCGCTGCGGGCGAGGTGAGGGGCGGAGGTGAGGGGCGAGGTTCCCAGCAGG
ATGCCCGGCTCTGCAGGAAGCTGAAGTGAGAGGCCGGAGAGGGCCCAGCCGCCGGGC
AGGATGACCAAGGCCGGCTGTTCCGGCTGTGGCTGGTGTGGTGTTCATGATCCT
GCTGATCATCGTGTACTGGACAGCGCAGGCCGCGCACTTCTACTGCACACGTCTTCT
CTAGGCCGACACGGGCCGCGCTGCCACGCCGGGCGACAGGGACAGGGAGCTCACG
GCCGACTCCGATGTCGACGAGTTCTGGACAAGTTCTCAGTGTGGCGTAAGCAGAGCGA
CCTTCCCAGAAAGGAGACGGAGCAGCCGCTGCCGGGAGCATGGAGGAGAGCGTGGAG
GCTACGACTGGTCCCCCGCGACGCCCGCGCAGCCCAGACCAGGGCGGCAGCAGGCGGAG
CGGAGGAGCGTGCTGCGGGCTTCTGCCAACCTCCAGCCTGGCTCCCCACCAAGGAGCG
CGCATTGACGACATCCCCAACTCGGAGCTGAGCCACCTGATGTTGGACGACGGCACGGG
CCATCTACTGCTACGTGCCAAGGTGGCTGCACCAACTGGAAGCGCGTGTGATGATCGTGT
AGCGGAAGCCTGCTGCACCGCGGTGCGCCCTACCGCGACCCGCTGCGCATCCCGCGAGCA
CGTGCACAACGCCAGCGCGCACCTGACCTCAACAAGTTCTGGCGCCGCTACGGGAAGCTCT
CCCGCCACCTCATGAAGGTCAAGCTCAAGAAGTACACCAAGTTCTCTCGTGCACGGG
TTCGTGCGCCTGATCTCCGCCTCCGCAGCAAGTTCGAGCTGGAGAACGAGGAGTTCTACCG
CAAGTTCGCCGTGCCCATGCTGCGGCTGTACGCCAACACACCAGCCTGCCCGCTGGCGC
GCGAGGCCTTCCCGCCTCAAGGTGTCTCGCCAACCTCATCCAGTACCTGCTGGAC
CCGCACACGGAGAAGCTGGGCCCTCAACGAGCACTGGCGCAGGTGTACCGCCTTGCCA
CCCGTGCAGATCGACTACGACTTCGTGGGAAGCTGGAGACTCTGGACGAGGACGCCGCG
AGCTGCTGCAGCTACTCCAGGTGGACCGGCAGCTCCGCTCCCCCGAGCTACCGAACAGG
ACCGCCAGCAGCTGGAGGAGGACTGGTTCGCAAGATCCCCCTGGCCTGGAGGCAGCAGCT
GTATAAAACTCTACGAGGCCGACTTTGTTCTCTCGCTACCCCAAGCCGAAAACCTCCTCC
GAGACTGAAAGCTTCGCGTTGCTTTCTCGCGTGCCTGGAACCTGACGCACGCGCACTCC
AGTTTTTTATGACCTACGATTGCAATCTGGCTTCTGTTCACTCCACTGCCTCTATCC
ATTGAGTACTGTATCGATATTGTTTTAAGATTAATATTCAGGTATTAATACGA

FIGURE 194

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56112
<subunit 1 of 1, 414 aa, 1 stop
<MW: 48414, pI: 9.54, NX(S/T): 4
MTKARLFRLWLVLGSVFMILLIIVYWDSAGAAHFYLHTSFSRPHTGPPLPTPGPDRDRELTA
DSDVDEFFLDKFLSAGVKQSDLPRKETEQPPAPGSMEESVRGYDWSPRDARRSPDQGRQQAER
RSVLRGFCANSILAFPTKERAFDDIPNSELSHLIVDDRHGAIYCYVPKVACTNWKRVMIVLS
GSLLHRGAPYRDPLRIPREHVHNASAHLTFNKFWRRYGKLSRHLMKVKLKKYTKFLFVRDPF
VRLISAFRSKFELENEEFYRKFAVPMLRLYANHTSLPASAREAFRAGLKVSFANFIQYLLDP
HTEKLAPFNEHWRQVYRLCHPCQIDYDFVGKLETLDDEAAQLLQLLQVDRQLRFPPSYRNRT
ASSWEEDWFAKIPLAWRQQLYKLYEADFVLFGYPK PENLLRD
```

Important features:

Signal peptide:

amino acids 1-31

N-glycosylation sites.

amino acids 134-137, 209-212, 280-283 and 370-373

TNFR/NGFR family cysteine-rich region protein

amino acids 329-332

FIGURE 195

TCGGGCCAGAATTCGGCACGAGGCCGGCACGAGGGCGACGGCCTCACGGGCTTGAGGTGA
AAGAGGCCAGAGTAGAGAGAGAGAGACCGACGTACACGG**G**CTACGGAACCGCCT
ATGCCGGGAAGGTGGTGGTGTGACCGGGGGCGGCCGCGCATCGGAGCTGGATCGTGC
GCCTTCGTGAACAGCGGGGCCGAGTGGTTATCTGCGACAAGGATGAGTCTGGGGCCGGC
CCTGGAGCAGGAGCTCCCTGGAGCTGTCTTATCCTCTGTGATGTGACTCAGGAAGATGATG
TGAAGACCCTGGTTCTGAGACCATCCGCCGATTGGCCGCCGGATTGTGTTGTCACAAAC
GCTGGCCACCACCCACCCCCACAGAGGCCCTGAGGAGACCTCTGCCAGGGATTCCGCCAGCT
GCTGGAGCTAACCTACTGGGACGTACACCTTGACCAAGCTGCCCTCCCTACCTGCCGA
AGAGTCAAGGGATGTCATCAACATCTCCAGCCTGGTGGGGCAATGCCAGGCCAGGCA
GTTCCCTATGTGGCCACCAAGGGGGCAGTAACAGCCATGCCAAAGCTTGCCCTGGATGA
AAGTCCATATGGTGTCCGAGTCAACTGTATCTCCCCAGGAAACATCTGGACCCGCTGTGGG
AGGAGCTGGCAGCCTTAATGCCAGACCCCTAGGGCCACAATCCGAGAGGGCATGCTGGCCAG
CCACTGGGCCGCATGGGCCAGCCGCTGAGGTGGGGCTGCCAGTGTCCCTGCCCTCCGA
AGCCAACCTCTGCACGGCATTGAACTGCTCGTGACGGGGGTGCAGAGCTGGGTACGGGT
GCAAGGCCAGTCGGAGCACCCCGTGGACGCCCGATATCCCTCCT**G**ATTCTCTCATTT
CTACTTGGGGCCCCCTTCCTAGGACTCTCCACCCAAACTCCAACCTGTATCAGATGCAGC
CCCCAAGCCCTAGACTCTAACGCCAGTTAGCAAGGTGCCGGTCACCCCTGCAGGTTCCCAT
AAAAACGATTTGCAGCC

FIGURE 196

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56045
<subunit 1 of 1, 270 aa, 1 stop
<MW: 28317, pI: 6.00, NX(S/T): 1
MATGTRYAGKVVVVTGGGRGIGAGIVRAFVNNSGARVVICDKDESGGRALEQELPGAVFILCD
VTQEDDVKTLVSETIRRFGRLLDCVVNNAGHHPPPQRPEETSAQGFRQLLELNLLGTYTLKL
ALPYLRKSQGNVINISSLVGAIGQAQAVPYVATKGAVTAMTKALALDES PYGVRVNCISPGN
IWTPLWEELAALMPDPRATIREGMLAQPLGRMGQPAEVGAAAVFLASEANFCTGIELLVTGG
AELGYGCKASRSTPVDAPDIPS
```

Important features:

N-glycosylation site.

amino acids 138-141

Short-chain alcohol dehydrogenase family protein

amino acids 10-22, 81-91, 134-171 and 176-185

FIGURE 197

AGGCGGGCAGCAGCTGCAGGCTGACCTTGCAGCTTGGCGGAATGGACTGGCCTCACAAACCTG
CTGTTTCTTCTTACCATTTCCATCTTCCTGGGCTGGGCCAGCCCAGGAGCCCCAAAGCAA
GAGGAAGGGCAAGGGCGGCTGGGCCCTGGCCCTGGCCCTCACCAGGTGCCACTGGACC
TGGTGTCACTGGATGAAACCGTATGCCGCATGGAGGAGTATGAGAGGAACATCGAGGAGATG
GTGGCCCAGCTGAGGAACAGCTCAGAGCTGGCCCAGAGAAAGTGTGAGGTCAACTTGCAGCT
GTGGATGTCCAACAAGAGGAGCCTGTCTCCCTGGGCTACAGCATCAACCACGACCCCCAGCC
GTATCCCCGTGGACCTGCCGGAGGCACGGTGCCTGTGTCTGGCTGTGAACCCCTTCACC
ATGCAGGAGGACCGCAGCATGGTGAGCGTGCCTGTTAGCCAGGTTCCGTGCGCCGCCG
CCTCTGCCGCCACCGCCCCGCACAGGGCTTGCCGCCAGCGCGAGTCATGGAGACCATCG
CTGTGGCTGCACCTGCATCTTCTGAATCACCTGGCCCAGAAGCCAGGCCAGCAGCCCCGAGA
CCATCCTCCTGCACCTTGTGCCAAGAAAGGCCTATGAAAAGTAAACACTGACTTTGAAA
GCAAG

FIGURE 198

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59294
<subunit 1 of 1, 180 aa, 1 stop
<MW: 20437, pI: 9.58, NX(S/T): 1
MDWPHNLLFLLTISIFLGLGQPRSPKSKRKGQGRPGPLAPGPHQVPLDLVSRMKPYARMEY
ERNIEEMVAQLRNSSELAQRKCEVNLQLWMSNKRSLSPWGYSINHDPSRIPVVDLPEARCLCL
GCVNPFTMQEDRSMVSVPVFSQVPVRRRLCPPPRTGPCRQRAVMETIAVGCTCIF
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 75-78

Homologous region to IL-17

amino acids 96-180.

FIGURE 199

GCAGCGCCAGCGTAGGCGGGTGGCCCTGCGTCTCCGCTTCAAGAAAACCGGGCGGG
CGAGCGAGGCTCGGGCGGCGCTGCCCTTCCCCACACTCCCGCCGAGAAGCCTCGCTCG
GCGCCCAACATGCGGGTGGCGCTCGGGCCCGAGCTAACGGCGCTCTGGCCGCTGGAT
CGCGGCTGTGGCGGCGACGGCAGGGCCCGAGGAGGGCGCGCTGCCGCGAGCAGAGCCGG
TCCAGCCCATGACCGCCTCCAACGGACGCTGGTGATGGAGGGCGAGTGGATGCTGAAATT
TACGCCCCATGGTGTCCATCCTGCCAGCAGACTGATTAGAATGGGAGGCTTGCAGAAAGAA
TGGTGAAATACCTCAGATCAGTGTGGGAACGTAGATGTCATTCAAGAACAGGTTGAGTG
GCCGCTTCCTTGTCAACCCTCTCCAGCATTTTCATGCAAAGGATGGATATTCCGCCGT
TATCGTGGCCCAGGAATCTCGAAGACCTGCAGAATTATATCTTAGAGAAGAAATGGCAATC
AGTCGAGCCTCTGACTGGCTGGAATCCCCAGCTCTCAACGATGCTGGAATGGCTGGTC
TTTTAGCATCTGGCAAGATATGGCATCTTACAACATTACAGTGACTCTGGAATT
CCTGCTTGGTGTCTTATGTGTTTTCGTATAGCCACCTGGTTTGGCCTTTATGGG
TCTGGTCTTGGTGGTAATATCAGAATGTTCTATGTGCCACTTCAAGGCATTATCTGAGC
GTTCTGAGCAGAATGGAGATCAGAGGAGGCTCATAGAGCTGAACAGTTGCAGGATGCCAG
GAGGAAAAAGATGATTCAAATGAAGAAGAAAACAAAGACAGCCTGTAGATGATGAAGAAGA
GAAAGAAGATCTGGCGATGAGGATGAAGCAGAGGAAGAGAGGAGGACAACCTGGCTG
CTGGTGTGGATGAGGAGAGAAGTGAGGCCATGATCAGGGGCCCCCAGGAGAGGACGGTGTG
ACCCGGGAGGAAGTAGAGCCTGAGGAGGCTGAAGAAGGCATCTGAGCAACCTGCCAGC
TGACACAGAGGTGGTGGAAAGACTCCTTGAGGCAGCGTAAAGTCAGCATGCTGACAAGGGAC
TGTAGATTAATGATGCGTTTCAAGAATACACAAAACATATGTCAGCTCCCTTGG
CCTGCAGTTGTACCAAATCTTAATTTCCTGAATGAGCAAGCTCTCTAAAAGATGCT
CTCTAGTCATTGGTCTCATGGCAGTAAGCCTCATGTATAACTAAGGAGAGTCTCCAGGTGT
GACAATCAGGATATAGAAAAACAAACGTAGTGTGGATCTGTTGGAGACTGGGATGGGAA
CAAGTTCATTACTTAGGGGTAGAGAGCTCGACCAGAGGAGGCCATTCCCAGTCCTAATC
AGCACCTTCCAGAGACAAGGCTGCAGGCCCTGTGAAATGAAAGCCAAGCAGGAGCCTTGGCT
CCTGAGCATTCCAAAGTGTAAAGCTAGAAGCCTTGACATCCTTCTTGTAAAGTATTAT
TTTGTCAAATTGCAAGGAAACATCAGGCACCACAGTGATGAAATCTTCACAGCTAGAA
ATTGAAAGGGCTTGGGTAGAGAGCAGCTCAGAAGTCATCCCAGCCCTCTGAATCTCCTG
TGCTATGTTTATTCTACCTTAATTTCAGCATTCCACCATGGCATTAGGCTCT
CCACACTCTTCACTATTCTCTGGTCAGAGGACTCCAATAACAGCAGGTTACATGAAC
TGTGTTGTTCACTCTGACCTAAGGGTTAGATAATCAGTAACCATAACCCCTGAAGCTGT
GACTGCCAAACATCTCAAATGAAATGTTGGCCATCAGAGACTCAAAAGGAAGTAAGGATT
TTACAAGACAGATTAATGTTTGTCCAAATATAGTTGTTGATTGTTT
AAGTTTCTAAGCAATATTTCAAGCCAGAAGTCCTCTAAGTCTGCCAGTACAAGGTAGT
CTTGTGAAGAAAAGTTGAATACTGTTTGTCTCATCTCAAGGGGTTCCCTGGGTCTTGAAC
TACTTTAATAATAACTAAAAACCACTCTGATTTCCTCAGTGATGTGTTGGTAAAGA
GAATTAATGAACCTCCAGTACCTGAAAGTGAAGATTGATTGTTCCATCTCTGTAATC
TTCCAAGAATTATATCTTGTAAATCTCAATACTCAATCTACTGTAAGTACCCAGGGAG
GCTAATTCTT

FIGURE 200

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56433
<subunit 1 of 1, 349 aa, 1 stop
<MW: 38952, pI: 4.34, NX(S/T): 1
MAGGRCGPQLTALLAAWIAAVAAATAGPEEAALPPEQSRVQPMTASNWTLVMEGEMLKFYAP
WCPSCQQTDSEWEAFAKNGEILQISVGKVDVIQEPGLSGRFFVTLPAFFHAKDGIFRRYRG
PGIFEDLQNYILEKKWQSVEPLTGWKSPASLTMSGMAGLFSISGKIWHLHNYFTVTLGIPAW
CSYVFFVIATLVFGLFMGLVLVVISECFYVPLPRHLSERSEQNRRSEEAHRAEQLQDAEEEK
DDSNNEENKDSLVDDEEEKEDLGDEDEAEEEEEDNLAAGVDEERSEANDQGPPGEDGVTR
EVEPEEAESEGISEQPCPADTEVVVEDSLRQRKSQHADKGL
```

Important features:

Signal peptide:

amino acids 1-22

Transmembrane domain:

amino acids 191-211

N-glycosylation site.

amino acids 46-49

Thioredoxin family proteins. (homologous region to disulfide isomerase)

amino acids 56-72

Flavodoxin proteins

amino acids 173-187

FIGURE 201

ATCTGGTTGAAC TACTTAAGCTTAATTGTTAAACTCCGGTAAGTACCTAGCCCACATGATT
TGACTCAGAGATTCTCTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAAGAGCTCC
CAAATGCTATATCTATTCAAGAACATGGAATATCATCCTGATTTAGAAAAT
TTGGATGAAGATGGATATACTCAATTACACTTCGACTCTCAAAGCAATACCAGGATAGCTGT
TGTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTGGCGCCTCATTGCTGTAATTGG
GAATCCTATGCTGGTAATACTGGTGTAGCTGTGCTGGTACCATGGGGTTCTTCC
AGCCCTTGTCTCCTAATTGGATTATATGAGAAGAGCTGTTATCTATTCAAGCATGTC
AAATTCTGGGATGGAAGTAAAGACAATGCTGGCAACTGGGCTCTAATCTCTAAAGATAG
ACAGCTCAAATGAATTGGGATTATAGTAAACAAGTGTCTCCAACCTGATAATTCA
TGGATAGGCCTTCTGGCCCCAGACTGAGGTACCATGGCTCTGGGAGGGATGGATCAACATT
CTCTCTAACTTATTTCAGATCAGAACACAGCTACCCAGAAAACCATCTCCAAATTGTG
TATGGATTACGTGTCAGTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTGT
GAGAAGAAGTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTA
AGGAGGACAGAAAACAGAACAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATG
TTTAGAGAGCTGGCAACTGTAATCTAACCAAGAAATTGAAGGGAGAGGCTGTGATTCT
GTATTGTCGACCTACAGGTAGGCTAGTATTATTTCTAGTTAGTAGATCCCTAGACATGG
AATCAGGGCAGCCAAGCTGAGTTTATTATTTTATTATTTTATTATTTTGTGAGATAGGGTCT
CACTTGTACCCAGGCTGGAGTCAGTGGCACAATCTGACTCACTGCAGCTATCTCGC
CTCAGCCCCCTCAAGTAGCTGGACTACAGGTGCATGCCACCATGCCAGGCTAATTGGTG
TTTTTGTAGAGACTGGGTTTGCCTGTTGACCAAGCTGGCTCTAACTCCTGGCTTAAG
TGATCTGCCCGCTTGGCTCCAAAGTGCCTGGGATTACAGATGTGAGGCCACCACCTGGC
CCCAAGCTTGAATTTCATTCTGCCATTGACTTGGCATTTACCTTGGTAAGCCATAAGCGA
ATCTTAATTCTGGCTCTATCAGAGTTGTTCATGCTCAACAATGCCATTGAAGTGCACGGT
GTGTTGCCACGATTGACCCCTCAACTCTAGCAGTATATCAGTTAGTAACTGAGGGTGAAAT
ATATTCTGAATAGCTAAATGAAGAAATGGAAAAAAATCTCACCACAGTCAGAGCAATT
ATTATTTCATCAGTATGATCATAATTATGATTATCATTTAGTAAAGCAGGAACCTCTA
CTTTTTCTTATCAATTAAATAGCTCAGAGACTACATCTGCCATATCTCTAATAGAATCTT
TTTTTTTTTTTTTTGAGACAGAGTTCGCTTGTGCTGCCAGGCTGGAGTGCAACGG
CACGATCTGGCTCACCGAACCTCCGCCCTGGGTTCAAGCAATTCTCTGCCAGCCT
CCCAAGTAGCTGGGATTACAGTCAGGCACCACACCCGGCTAATTGTATTGGTAGT
AGAGACAGGGTTCTCATGTCGGTCAGGGTAGTCCGAACCTGACCTCAAGTGTCTGC
CTGCCTCGGCCTCCCAAGTGTGGATTACAGGCGTGAGCCACTGCACCCAGCCTAGAATCT
TGTATAATATGTAATTGTAAGGAAACTGCTCTCATAGGAAAGTTCTGCTTTAAATACA
AAAATACATAAAATACATAAAATCTGATGATGAATATAAAAAGTAACCAACCTCATTGGA
ACAAGTATTAACATTGGAATATGTTTATTAGTTGTGATGACTGTTTACAATT
ACCATTTTTCAGTAATTACTGTAAGGTTATTGGAATGAAACTATATTCTCATG
TGCTGATTGCTTATTCTCATACTTCCCAGTGGCTATTGTTATTCCTGAAATGGATA
TTCTGTATTACTAGGGAGGCATTACAGTCCTCTAATGTTGATTAATATGTGAAAGAAAT
TGTACCAATTAAATTATGCAGTTAAATGGATGATTATGTTATGTGGATTTCAT
TTCAATAAAAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 202

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912
<subunit 1 of 1, 201 aa, 1 stop
<MW: 22563, pI: 4.87, NX(S/T): 1
MEYHPDLENLDEDGYTQLHFDSQSNTRIAVVSEKGSCAASPPWRLIAVILGILCLVILVIAV
VLGTMGVLSSPCPPNIIYEKSCYLFMSILNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQ
VSSQPDNSFWIGLSRPQTEVPWLWEDGSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQL
CSVPSYSICEKKFSM
```

Important features:

Type II transmembrane domain:

amino acids 45-65

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 197-200

N-myristoylation sites.

amino acids 35-40 and 151-156

Homologous region to LDL receptor

amino acids 34-67 and 70-200.

FIGURE 203

FIGURE 204

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50921
<subunit 1 of 1, 693 aa, 1 stop
<MW: 77738, pI: 8.87, NX(S/T): 7
MTPQSLLQTTLFLLSLLFLVQGAHGRGHREDFRFCSQRNQTHRSSLHYKPTPDLRISIENSE
EALTVHAPFPAAHPASRSFPDPRGLYHFCLYWNRHAGRLHLLYGKRDFLSDKASSLLCFQH
QEEESLAQGPPLLATSVTWSWPQNISLPSAASFTFSFHSPPPHTAAHNASVDMCELKRDILQLL
SQFLKHPQKASRRPSAAPASQQLQSLESKLTSVRFMGDMVSFEEDRINATVWKLQPTAGLQD
LHIHSRQEEEQSEIMEYSVLLPRTLFQRTKGRSGAEKRLLLVDFSSQALFQDKNNSQVLGE
KVLGIVVQNTKVANLTEPVVLTFQHQQLQPKNVTLQCVFWVEDPTLSSPGHSSAGCETVRRE
TQTSCFCNHLYFAVLMVSSVEVAVHKHYLSSLSYVGCVVSALACLVTIAAYLCSRVPPLPC
RRKPRDYTIKVHMNLLAVFLLDTSFLLSEPVALTGSEAGCRASAIFLHFSLLTCLSWMGLE
GYNLYRLVVEVFGTYVPGYLLKLSAMGWGFPIFLVTVALVDVDNYGPIILAVHRTPEGVIY
PSMCWIRDSLVSYITNLGLFSLVFLFNAMMLATMVQILRLRPHTQKWSHVLTLLGLSLVLG
LPWALIFFSFASGTFQLVVLYLFSIITSFQGFLIFIWYWSMRLQARGGSPPLKSNSDSARLP
ISSGSTSSSRI
```

Important features:

Signal peptide:

amino acids 1-25

Putative transmembrane domains:

amino acids 382-398, 402-420, 445-468, 473-491, 519-537, 568-590
and 634-657

Microbodies C-terminal targeting signal.

amino acids 691-693

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 198-201 and 370-373

N-glycosylation sites.

amino acids 39-42, 148-151, 171-174, 234-237, 303-306, 324-327
and 341-344

G-protein coupled receptors family 2 proteins

amino acids 475-504

FIGURE 205

TGCCTGGCCTGCCTTGTCAACAATGCCGCTTACTCTGCTTCCAGGTTGCCCTGCCCTGCAGA
GGAAANCNTCGGGACTACACCNCAAGTGCACATGAACCTGCTGCCGTCTCCTGCTG
GACACGAGCTTCTGCTCAGCGNAGCCGGTGGCCCTGACAGGCTCTGAAGGCTGGCTGCCGA
GCCAGTGCCATCTCCTGCACTTCTCCTGCTCACCTGCCTTCCTGGATGGCCTCGAGGGG
TACAACCTCTACCGACTCGTGGTGGAGGTCTTGGCACCTATGTCCCTGGCTACCTACTCAA
GCTGAGCGCCATGGGCTGGGCTTCCCCATCTTCTGGTGACGCTGGTGGCCCTGGATG
TGGACAACATGGCCCCATCATCTTGGCTGTGCATAGGACTCCAGAGGGCGTCATCTACCC
TCCATGTGCTGGATCCGGGACTCCCTGGTCAGCTACATACCAACCTGGCCTTTAGCCT
GGTTCTGTTAACATGG

FIGURE 206

CGGACGCGTGGCGGACCGTGGCGGACCGCGTGGCGGACCGCGTGGCGGCTAGGAAAGAGGGCTCTAGGAAAAGTTGGATGGGATTATGTGAAACTACCCCTGCGATTCTCTGCTGCCAGAGCAGGCTCGCGCTTCCACCCAGTGCAGCCTCCCTGGCGGTGGTGAAGAGACACGGGAGTCGCTCTCACCCAGTCAGCCAAATGACCCCTTCGGCTTCTCGGGCTTCCTGCTGACATCTGCCCTGGCGGCCAGAGACAGGGACTCAGGGGAATCCTGAGTAGTAAATTCCAGTTTCCAGCAACAAGGAACAGAACGGAGTACAAGATCCTCAGCATGAGAGAATTATTACTGTGTCTACTAATGGAAGTATTACAGGCCAAGGTTCTCTCATACTTATCCAAGAAATACGGCTTGGTATGGAGATTAGTAGCAGTAGAGAAAATGTATGGATACAACCTACGTTGATGAAAGATTGGGCTTGAAGACCCAGAAGATGACATATGCAAGTATGATTGTAGAAGTTGAGGAACCCAGTGTGAACTATATTAGGGCGTGGTGGTTCTGGTACTGTACCAAGAAAACAGATTCTAAAGGAATCAAATTAGGATAAGATTGTATCTGATGAATATTTCCTCTGAACCAGGGTCTGCATCCACTACAACATTGTCATGCCACAATTACAGAACGGTGTGAGTCCTCAGTGTACCCCTTCAGCTTGGCACTGGACCTGCTTAATAATGCTATAACTGCCTTAGTACCTTGGAAAGACCTTATTGATATCTTGAACCAAGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACCTGGCAACTCTTGGCAAGGTTTGGAAAGAAATCCAGAGTGGTGGATCTGAACCTCTAACAGAGGAGGTAAAGATTATAACAGCTGCACACCTCGTAACCTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTCTGGCCAGGTTGTCTCCTGGTTAACCGCTGTGTTGGGAACGTGTGCTTGTGCTCCACAATTGCAATGAATGTCAATGTGTCCAAAGCAAAGTTACTAAAAAATACCACGAGGTCTTCAGTTGAGACCAAAGACCGGTGTCAAGGGATTGACAAATCACTCACCGACGTGGCCCTGGAGCACCAGTGCAGTGGCTTGTGCTTCAAGGACCTTCATCTCAGGATTACAGTCAGTGTGCTTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTGGCTTAAATAGATCACCAGCTAGTTCAAGGTACAGTACGTATTCCACTAGCTGGGTTCTGTATTCTAGTTCTTCGATACGGCTTAGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGTGAGCACCTGATTCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCTTAAATCGTATAAAATCTGATTTTTTTTTTTGCTCATATTACATATGTAACACCAGAACATTCTATGACTACAAACCTGTTTAAAGGAACATGTGCTATGAAATTAAACTTGTGTCATGTTGCTGAGTGGTAAAGAAATTAAATGTTGTTAAACATGGTGTGCTGAGTGGGATTTCTATTTCTTATTAAATTTCTGCCATTAGAAGAAGAGAACATACATTCATGGTTTGGAAAGAGATAAAACCTGAAAAGAGAGTGGCCTTATCTTCATTTATCGATAAGTCAGTTATTGTTTCATTGTGTACATTTTATATTCTCCTTTGACATTATAACTGTGCTTCTAATCTGTTAAATATATCTATTTTACAAAGGTATTAAATTCTTTATGACAACCTAGATCAACATTGTTAGCTTGGTAAAGAAATTCTTCAAAACACAATTGTTAGGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAAATACATGTATTCAATTCTCGTATGGTCTAGAGTTAGATTAATCTGCAATTAAACTGAATTGGAATAGAATTGGTAAGTTGCAAGACTTTTGAAAGATAATTAAATTATCATATCTCCATTCTGTATTGGAGATGAAAATAAAAAGCAACCTATGAAAGTAGACATTCAAGATCCAGCATTACTAACCTATTCTTTTGGGAAATCTGAGCCTAGCTCAGAAAAACATAAGCACCTTGAAAAGACTTGGCAGCTCCTGATAAAAGCGTGTGCTGAGTAGGAACACATCCTATTATGTTGATGTTGCTGGTTTATTATCTTAAACTCTGTCATACACTGTATAAAATACATGGATATTGTTATGTACAAGTAGTGTCTCTAACCACTGTCACATTGTAACCTGTGCAATTAAAGAAAATCAGTAAAATATTGCTTGTAAAATGCTTAAATATNGTCCTAGGTTATGTGGTGAACATTGTAATCAAATGTTGAATCATCAAATAAAGAATGTGGCTATTGGGGAGAAAATTAAAAAAAAAAAGGTTAGGGATAACAGGGTAATGCGGCC

FIGURE 207

MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERIITVSTNGSIHSPR
FPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPEDDICKYDFVEVEEPSDGTILGRWC
GSGTVPGKQISKGNQIRIRFVSDEYFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLL
NNAITAFSTLEDLIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY
SCTPRNFSVSIREELKRTDTIFWPGC₁₄LLVKRCGGNCACCLHNCNECQCVP₁₄SKVTKKYHEVLQ
LRPKTGVRGLHKSLTDVALEHHEECDCVCRG₁₄STGG

Signal sequence:

amino acids 1-14

100 200 300 400 500 600 700 800 900 1000

FIGURE 208

CCCATCTCAAGCTGATCTGGCACCTCTCATGCTCTGCTCTTCAACCAGACCTCTACATTCCATTGGAAAGA
AGACTAAAATGGTGTTCATGTCAGAGACAAATTCTTACATAATCCTAATTCC
AAACTCCTGGGGCTAGATGGTTCTAAACTCTGCCCTGTGATGTCCTGATGTCACCTGGATGTCACCTGGAAAGAACCATGTG
ATCGTGGACTGCACAGACAAGCATTGACAGAAATTCTGGAGGTATTCCCACGAACACCACGAACCTCACCC
ACCATTAAACCACATACCAAGACATCTCCCAGCGTCCTTCACAGACTGGACCATCTGGTAGAGATCGATTTCAGA
TGCAACTGTGTACCTATTCCACTGGGGTCAAAAAACAAACATGTGATCAAGAGGCTGCAGATTAAACCCAGAACG
TTTAGGGACTCACTTATTTAAATCCCTTACCTGGATGAAACCAAGCTACTAGAGATACCGCAGGGCCTCCG
CCTAGCTTACAGCTCTCAGCCTGAGGCCAACACATCTTTCCATCAGAAAAGAGAATCTAACAGAACTGGCC
AACATAGAAATACTTACCTGGGCCAAACTGTTATTATCGAAATCCTGTTATGTTCATATTCAGAGAAA
GATGCCCTCCTAAACTGACAAAGTTAAAGTGTCTCCCTGAAAGATAACAAATGTCACAGCGTCCCTACTGTT
TTGCCATCTACTTTAACAGAACTATATCTCTACAACAAACATGATTGCAAAATCCAAGAAGATGATTAAATAAC
CTCAACCAATTACAAATTCTGACCTAAGTGGAAATTGCCCTCGTTATAATGCCCAATTCTGTGCGCCG
TGTAAAAATAATTCTCCCTACAGATCCCTGAAATGCTTTGATGCGTGCAGAAATTAAAGTTTACGTCTA
CACAGTAACCTCTCAGCATGTGCCCAAGATGGTTAAGAACATCAACAAACTCCAGGAACGGATCTGTCC
AAAAACTCTGGCCAAGAAATTGGGGATGCTAAATTCTGATTTCTCCAGCCTCATCCAATTGGATCTG
TCTTCAAATTGAACTTCAGGTCTATCGTGCATCTATGAATCTACAAAGCATTCTTCACTGAAAAGCCTG
AAAATTCTGCGGATCAGAGGATATGCTTTAAAGAGTTGAAAAGCTTAACTCTCGCCATTACATAATCTCAA
AATCTTGAAGTTCTGATCTGGCACTAATTAAATTGCTAACCTCAGGAGATTCAAGTGAAGTGGCTCTGCTCAAAT
CTGAAAGTCAGATCTTCAGTGAATAAAATATCACCTTCAGGAGATTCAAGTGAAGTGGCTCTGCTCAAAT
GCCAGAACTCTGTAGAAAGTTATGACCCCCAGGTTCTGGAAACAAATTACATTATTTCAGATATGATAAGTATGCA
AGGAGTTGAGATTCAAAACAAAGAGGTTCTTCATGTCGTTAATGAAAGCTGCTACAAGTATGGCAGACC
TTGGATCTAAGTAAAATAGTATATTGTCAGGCTCTGATTTTCAGCATTCTTCTCCTCAAATGCCCTG
AATCTGTCAGGAAATCTCAATTAGCCAATTCTTCAACAGCATTGAAAGAGCTTCACAAACTGGAAAGTTCTGGAT
GACTTCCTCAACACCAGGCTTGATTTACTCCATTCAACAGCATTGAAAGAGCTTCACAAACTGGAAAGTTCTGGAT
ATAAGCAGTAATAGCATTATTTCATCAGAAGGAATTACTCATATGCTAACCTTACCAAGAACCTAAAGGTT
CTGCAGAAACTGATGATGAACGACAATGACATCTCCCTCACCAGCAGGACCATGGAGAGTGTCTTCTAGA
ACTCTGGAATTCAAGAGGAATCACTTAGATGTTATGGAGAGAAGGTGATAACAGATACTTACAATTATTCAAG
AATCTGCTAAAATTAGAGGAATTAGACATCTCTAAATTCCCTAAGTTCTGCTCTGGAGTTTGATGGT
ATGCCTCAAATCTAAAGAATTCTCTTGGCCAAAATGGGCTCAAATCTTCAGTGGAAAGAACCTCAGTGT
CTAAAGAACCTGGAAACTTGGACCTCAGCCACAACCAACTGACCACTGTCCTGAGAGATTATCCAACCTGTTCC
AGAACGCCAAGAACTGATTCTTAAGAATAATCAAATCAGGAGCTGACGAAGTATTCTACAAAGATGCCCTC
CAGTTGCGATATCTGGATCTCAGCTCAAATAAAATCCAGATGATCCAAAGACCAGCTCCAGAAAATGTCCTC
AAACATCTGAAAGATGTTGCTTTGCATCATCAGGTTCTGTGCACCTGTGATGCTGTGGTTGTGGTGG
GTAAACCATACGGAGGTGACTATTCTTACCTGGCACAGATGTGACTTGTGTGGGCCAGGACACAAAGGGC
AAAAGTGTGATCTCCCTGGATCTGACACCTGTGAGTTAGATCTGACTAACCTGATTCTGTTCTACTTCCATA
TCTGTATCTCTTCTCATGGTGTGATGACAGCAAGTCACCTCTATTCTGGGATGTGTGGTATATTACCAT
TCTGTGTAAGGCAAGATAAAAGGGTATCAGCGTCTAATATCACCAGACTGTTGCTATGATGCTTTATTGTGTAT
GACACTAAAGACCCAGCTGTGACCGAGTGGGTTTGGCTGAGCTGGCCAAACTGGAAGACCAAGAGAGAAA
CATTTTAATTATGTCAGGAGAAAGGGACTGGTTACCAAGGGCAGCCAGTCTGGAAAACCTTCCAGAGCATA
CAGCTTAGCAAAAGACAGTGGTTGTGATGACAGACAAGTATGCAAAGACTGAAAATTAAAGATAGCATTTCAC
TTGTCCCCTCAGAGGCTCATGGATGAAAAGTTGATGATTATCTGATATTCTTGTGAGAAGGCCCTTCAGAAG
TCCAAGTTCTCCAGCTCCGGAAAAGGCTCTGTGGGAGTTCTGTGCTTGTGAGTGGCCAAACAAACCCGCAAGCTCAC
CCATAACTCTGGCAGTGTCTAAAGAACGCCCTGGCACAGACAATCATGGCCTATAGTCAGGTGTTCAAGGAA
ACGGTCTAGCCCTTCTTGCAAAACACAACGCCTAGTTACCAAGGAGAGGCCTGGC

FIGURE 209

MVFPMWTLKRQILILFNIILISKLLGARWFPKTLPCDVLDVPKNHVIVDCTDKHLTEIPGG
IPTNTTNLTLTINHIPDISPASFHRLDHLVEIDFRCNCVPIPLGSKNNMCIKRLQIKPRSFS
GLTYLKSLYLDGNQOLLEIPQGLPPSLQLLSLEANNIFSIRKENTELANIEILYLGQNCYYR
NPCYVSYIEKDAFLNLTKVLKVLQSLKDNNVTAVPTVLPSTLTELQYLYNNMIAKIQEDDFNNL
NQLQILDLSGNCPRCYNAPFPCAPCKNNPLQIPVNAFDALTELKVRLHSNSLQHVPPRWF
KNINKLQELDLSQNFLAKEIGDAKFLHFLPSLIQLDLSFNFELQVYRASMNLSQLAFSSLKSL
KILRIRGYVFKELKSFNLSPFHNLQNLQEVLDLGTNFIKIANLSMFQFKRLKVIDLSVNKIS
PSGDSSEVGFCNSARTSVESYEPQVLEQLHYFRYDKYARSCRFKNKEASFMSVNESCYKGQ
TLDLSKNSIFFVKSSDFQHLSFLKCLNLSGNLISQTLNGSEFQPLAELRYLDFSNRNLQD
STAFEELHKLEVLDISSLNHYFQSEGITHMLNFTKVLQKLMNDNDISSTSRTMESES
LRTLEFRGNHLDVLWREGDNRYLQLFKNLLKLEELDISKNSLSFLPSGVFDGMPPNLKNL
AKNGLKSFSWKKLQCLKNLETLDLSHNQLTTVPERLSNCSRSLKNLILKNNQIRSLTKYFLQ
DAFQLRYLDLSSNKIQMIQKTSFPENVLNNLKMLLHHNRFLCTCDAWFVWWVNHTEV
YLATDVTCVGPGAHKGQSVISLDLYTCELDLTNLILFSLSISVSLFLMVMMTASHLYFW
YIYHFCKAKIKGYQRLISPDCCYDAFIVYDTKDPAVTEWVLAELVAKLEDPREKHFNL
RDWLPGPVLENLSQSIQLSKTVFVMTDKYAKTENFKIAFYLHQRLMDEKVDVII
KPFQKSKFLQLRKRLCGSSVLEWPTNPQAHPYFWQCLKNALATDNHVAYSQVF
KETV

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 840-860

FIGURE 210

GGGTACCATCTGGCTGCTGCAAGTTACGGAAATGAAAAATTAGAACAAACAGAAACATGGAAAACATGTTCCCTTC
AGTCGTCATGTCACCTGCATTTCTGCTAATATCTGGTCTGTGAGTTATGCGCCGAAGAAAATTTCTA
GAAGCTATCCTTGTGATGAGAAAAGCAAATGACTCAGTTATGCCAGAGTGCAGCAATCGTCGACTACAGGAAG
TTCCCCAAACGGTGGCAAATATGTGACAGAACTAGACCTGTGATAATTTCATCACACACATAACGAATGAAT
CATTTCAGGGCTGCAAATCTCACTAAAATAATCTAAACACACCCAAATGTACAGCACAGAACGGAAATC
CCGGTATACAATCAAATGGCTGAATATCACAGACGGGCATTCCTCAACCTAAAACCTAAGGGAGTTACTGC
TTGAAGACAACCAGTTACCCCAAATACCCCTGGTTGCCAGAGTCTTGAGAAGACTTAGTCTAATTCAAACAA
ATATATACAACATAACTAAAGAGGGCATTCAGACTTAAACATTGAAAATCTCTATTGGCCTGAACTGCT
ATTTAACAAAGTTGCGAGAAAACTAACATAGAACAGATGGAGTATTGAAACGCTGACAATTGGAGTTGCTAT
CACTATCTTCATTCACACGTGCCACCCAACTGCCAGCTCCCTACGCAAACATTTCAGCAACA
CCCAGATCAAATACATTAGTGAAGAAGATTCAAGGGATTGATAAAATTAAACATTACTAGATTAAAGCGGGAACT
GTCCGAGGTGCTCAATGCCCAATTCCATGCGTGCCTGTGATGGTGGTGCCTCAATTAAATAGATGTTTG
CTTTTCAAAACTTGACCCAACTCGATAACCTAAACCTCTAGCACCTCCCTCAGGAAGATTAAAGCTGCCTGGT
TTAAAAATATGCCTCATCTGAAGGTGCTGGATCTGAATTCAACTATTAGTGGAGAAATAGTCTCTGGGCAT
TTTAACGATGCTGCCCGCTAGAAATACTGACTTGTCTTTAACTATATAAAGGGAGTTACCCACAGCATA
TTAATATTCCAGAAACTCTCTAAACTTTGTCTCACGGCATTGCAATTAAAGAGTTATGTGTTCCAGGAAC
TCAGAGAAGATGATTCCAGCCCTGATGCAAGCTCCAAACTTATGACTATCAACTTGGTATTAAATTAA
AGCAAATCGATTTCAAACCTTCCAAATCTGAAATTATTACTTGTCAAGAAACAGAATATCAC
CGTTGGTAAAGATAACCGGAGAGTTGCAAATAGTCTCTTCAACGTCATATCCGGAAACGACGCTCAA
CAGATTGAGTTGACCCACATTGAACTTTATCATTTCACCCGCTTTAATAAAGCCACAATGTGCTGCTT
ATGAAAAGCCTTAGATTAAAGCCTCAACAGTATTCTTCATTGGGCCAAACCAATTGAAAATCTCCTGACA
TTGCCCTGTTAAATCTGCTGCAAATAGCAATGCTCAACTGTTAAGTGGAACTGAATTTCAGCCATTCTCATG
TCAAATATTGGATTGACAACAAATAGACTAGACTTGTATAATGCTAGTGTCTTACTGAATTGTCGACTTGG
AAGTTCTAGATCTCAGCTATAATTACACTATTGAAATAGCAGGGCTAACACATCATCTAGAAATTATTCAA
ATTTCACAAATCTAAAGTTAAACATTGAGCCACAAACAAACATTAAACAGATAAGTATAACCTGGAAA
GCAAGTCCTGGTAAATTGTTTCACTGTTGCAATGCTGAGTGTGAACTGAATTTCAGCCATTCTCATG
TCTCCATTTCAAAGGTCTCAAGAATCTGACACGTCGGATTATCCCTTAATAGGCTGAAGCACATCCCAAATG
AAGCATTCTTAATTGCCAGCGAGTCTACTGAACATATAAATGATAATATGTTAAAGTTTAACTGGA
CATTACTCCAGCAGTTCTCGTCTGAGTTGACTTACGTGAAACAAACTACTCTTTAACTGATAGCC
TATCTGACTTACATCTTCCCTCGGACACTGCTGAGTCATAACAGGATTCCACCTACCCCTGGCTTTC
TTCTGAAAGTCAGTAGTCTGAAAGCACCTCGATTAAAGTCCAATCTGCTAAAACAATCAACAAATCCGACTT
AAACTAAGACCACCACCAATTATCTATGTTGAAACTACACGGAAACCCCTTGAATGCACCTGTGACATTGGAG
ATTCCGAAAGATGGATGGATGAAACATCTGAATGTAAAATTCCAGACTGGTAGATGTCATTGTGCCAGTCTG
GGGATCAAAGAGGGAAAGAGTATTGAGTCTGGAGCTAACAAACTTGTGTTCAAGATGTCACTGCACTGATATT
TTTCTCACGTTTATCACCACCATGGTTATGTTGGCTGCCCTGGCTACCATTGTTACTGGATGTT
GGTTTATATATAATGTGTTAGCTAAGGTAAGGCTACAGGTCTCTTCCACATCCAAACTTCTATGATG
CTTACATTCTTATGACACCAAAGATGCCCTGTTACTGACTGGTGAATAATGAGCTGCGTACCCACCTGAAG
AGAGCCGAGACAAAACGTTCTCTTGTCTAGAGGGAGAGGGATTGGGACCCGGATTGCCATCATGACAACC
TCATGCGAGGACATCAACCAAAGCAAGAAAACAGTATTGTTAAACAAAAAATATGCAAAGCTGAACTTTA
AAACAGCTTTTACTGGCTTGTGAGGCTAAATGGATGAGAACATGGATGTGATTATATTATCCTGCTGGAGC
CAGTGTACAGCATTCTCAGTATTGAGGCTACGGCAGCGGATCTGAAGACTCCATCCTCAGTGGCCTGACA
ACCCGAAGGCAGAAGGCTGTTGGCAAACCTCTGAGAAATGTGGTCTTGACTGAAAATGATTACGGTATAACA
ATATGTATGTCGATTCCATTAAAGCAATACTAACTGACGTTAAGTCACTGATTTCGGCCATAATAAAGATGCAAAG
GAATGACATTCTGATTAGTTATGCTATGTAACAAATTATCCAAAACCTTAGTGGTTAAAACAACACA
TTTGCTGCCACAGTTTGAGGCTCAGGAGTCCAGGCCAGCATAAACTGGGTCCTCTGCTCAGGGTGTCTCAG
AGGCTCAATGTAGGTGTTCACCGAGACATAGGCATCACTGGGTCACACTCATGTGGTTGTTCTGGATTCA
ATTCCCTGGGCTATTGGCCAAAGGCTATACTCATGTAAGGCACTGCGAGGCTCTCCACAGGAGCTTGGCT
ATCAGAGCTAGCAAAAAGAGAGGGTTGCTAGCAAGATGAACTGACAATCTTGTGAAATGAAACAAAAAGTGAT
ATCTCATCACTTGGCCATATTCTATTGTTAGTAAAGTAAACACAGGTCCCACCGAGCTCCATGGGAGTGACCA
TCAGTCCAGGGAAACAGCTGAAGACCAAGATGGTAGCTCTGATTGCTTCAGTTGGCATCAACTATTCCCT
TGACTGCTGTCCTGGGATGGCTGCTATCTGATGATAGATTGAAATATCAGGAGGAGGGACTCTGTTGGACC
ATCTTAGCAGTTGACCTAACACATCTTCTTCAATATCTAAGAACCTTGCCTGACTAATGGTCTAATA
TTAAGCTGTTGTTATATTATCATATATCTATGGCTACATGGTTATATTATGCTGTTGCGTCCGGTTTAT
TTACAGTTGCTTTACAAATATTGCTGAACTTGAACATTGACTTCAAGGTTAGATGCCATTAAAGAACTGAGATGG
ATAGTTTAAAGCATTTTACTCTTACCAATTAAAAGTATGCAAGCTAAATTGAAAGCTTGGTCTATA
TTGTTAATTGCCATTGCTGAAATCTAAATGAATAAAATGTTCAATTACAAAAA

FIGURE 211

MENMFLOSSMLTCIFLLISGSCELCAEENFSRSYPCDEKKQNDSVIAECASNRLQEVQPTVG
KYVTELDLSDNFITHITNESFQGLQNLTKINLNHNPNVHQNGNPGIQSNGLNITDGAFLNL
KNLRELLLEDNQLPQIPSGLPESLTELSTLQNNIYNITKEGISRLINLKNLYLAWNCFNKV
CEKTNIEDGVFETLTNLELLSLSFNSLSHVPPKLPSSLRKLFLSNTQIKYISEEDFKGLINL
TLLDLSGNCPRCFNAPFPCVPCDGGASINIDRFAFQNLTQLRYLNLSSTSLRKINAAWFKNM
PHLKVLDFNYLVGEIVSGAFLTMLPRLEILDLSFNYIKGSYPQHINISRNFSKLLSLRAL
HLRGYVFQELREDDFQPLMQLPNLSTINLGINFIQIDFKLFQNFNSNLEIIYLSENRISPLV
KDTRQSYANSSSFQRHIRKRRSTDFFDPHSNFYHFTRPLIKPQCAAYGKALDLSLNSIFFI
GPNQFENLPDIACLNLSANSNAQVLSGTEFSAIIPHVKYLDLTNNRLDFDNASALTELDLEV
LDLSYNSHYFRIAGVTHLEFIQFTNLKVLNLSHNNIYTLDKYNLESKSLVELVFSGNRL
DILWNDDDNRYISIFKGLKNLTRLDSLNRLKHIPNEAFLNLPASLTELHINDNMLKFFNWT
LLQQFPRLELLDLRGNKLLFTDSDFSSLRTLLLSHNRISHLPSGFLSEVSSLKHLDLS
SNLLKTINKSALETKTTKLSMLELHGNPFECTCDIGDFRRWMDEHLNVKIPRLVDVICASP
GDQRGKSIVSLELTTCVSDVTAVILFFFITTMVMLAALAHHLFYWDVWFIVNVCLAKVK
GYRSLSTSQTFYDAYISYDTKDASVTDWVINELRYHLEESRDKNVLLCLEERDWDPLAIID
NLMQSINQSKKTVFVLTKKYAKSWNFKTAFYLALQRLMDENMDVIIIFILLEPVLQHSQYLR
RQRICKSSILQWPDPNPKAEGLFWQTLRNVVLTENDSRYNNMYVDSIKQY

Signal sequence:

amino acids 1-26

Transmembrane domain:

amino acids 826-848

FIGURE 212

400 399 398 397 396 395 394 393 392 391 390 389 388 387 386 385 384 383 382 381 380 379 378 377 376 375 374 373 372 371 370 369 368 367 366 365 364 363 362 361 360 359 358 357 356 355 354 353 352 351 350 349 348 347 346 345 344 343 342 341 340 339 338 337 336 335 334 333 332 331 330 329 328 327 326 325 324 323 322 321 320 319 318 317 316 315 314 313 312 311 310 309 308 307 306 305 304 303 302 301 300 299 298 297 296 295 294 293 292 291 290 289 288 287 286 285 284 283 282 281 280 279 278 277 276 275 274 273 272 271 270 269 268 267 266 265 264 263 262 261 260 259 258 257 256 255 254 253 252 251 250 249 248 247 246 245 244 243 242 241 240 239 238 237 236 235 234 233 232 231 230 229 228 227 226 225 224 223 222 221 220 219 218 217 216 215 214 213 212 211 210 209 208 207 206 205 204 203 202 201 200 199 198 197 196 195 194 193 192 191 190 189 188 187 186 185 184 183 182 181 180 179 178 177 176 175 174 173 172 171 170 169 168 167 166 165 164 163 162 161 160 159 158 157 156 155 154 153 152 151 150 149 148 147 146 145 144 143 142 141 140 139 138 137 136 135 134 133 132 131 130 129 128 127 126 125 124 123 122 121 120 119 118 117 116 115 114 113 112 111 110 109 108 107 106 105 104 103 102 101 100 99 98 97 96 95 94 93 92 91 90 89 88 87 86 85 84 83 82 81 80 79 78 77 76 75 74 73 72 71 70 69 68 67 66 65 64 63 62 61 60 59 58 57 56 55 54 53 52 51 50 49 48 47 46 45 44 43 42 41 40 39 38 37 36 35 34 33 32 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1

CCAGGTCCAACTCACCTCGGTTCTATCGATTGAATTCCCCGGGATCCTCTAGAGATCCCT
CGACCTCGACCCACCGTCCGCCAAGCTGGCCCTGCACGGCTGAAGGGAGGCTCCTGTGGA
CAGGCCAGGCAGGTGGCCTCAGGAGGTGCCAGGGCGGCAGTGGCCTGAGGCCAGCAGC
AAGGGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCGG
TCCAGCAGCATCAGCAGCCCCCAGGACCGGGAGGCACAGGTGGCCCCACCACCCGGAGGA
GCAGCTCCTGCCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAA
GGCCACCCCGCCTGGAGGCACAGGCATGAGGGCTCTCAGGAGGTGCTGCTGATGTGGCTT
CTGGTGTGGCAGTGGCGGACAGAGCACGCCAACGGCCAGGAGGTGCTGAGGGTGTGCTGT
CCGGGCTCACGGGACCCCTGTCTCGAGTCGTTGAGCGTGTGACCAGCCCTCCTCA
CCACCTGCGACGGGCACCGGCCGTCAGCACCTACCGAACCATCTATAGGACCGCCTACCGC
CGCAGCCCTGGCTGGCCCTGCCAGGCCTCGCTACCGTGTGCCCGCTGGAAGAGGAC
CAGCGGGCTTCCTGGGCGCTGTGGAGCAGCAATATGCCAGGCCATGCCGAACGGAGGGA
GCTGTGCTCAGCCTGGCGCTGCCGCTGCCCTGCAGGATGGCGGGGTGACACTGCCAGTCA
GATGTGGATGAATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCATCAACACCGCCGG
CAGTTACTGGTGCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGC
CCAAGGGAGGGCCCCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAA
GAAGTGCAGAGGCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGGC
CCCACTGCACAGCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCGGCAGCCTCC
TGGTGCACTCCTCCAGCAGCTGGCGCATCGACTCCCTGAGCGAGCAGATTCCCTCCTG
GAGGAGCAGCTGGGTCCTGCTCCTGCAAGAAAGACTCGTGACTGCCAGCGCCCCAGGCTG
GAUTGAGCCCCCTACGCCGCCCTGCAGCCCCATGCCCTGCCAACATGCTGGGGTCCAG
AAGCCACCTGGGGTGAUTGAGCGGAAGGCCAGGCAGGGCCTCCTCTCTCCTCCCC
TTCCTCGGGAGGCTCCCCAGACCCCTGGCATGGATGGGCTGGGATCTTCTCTGTGAATCCAC
CCCTGGCTACCCCCACCCCTGGCTACCCCAACGGCATCCCAAGGCCAGGTGGGCCCTCAGCTG
AGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGGACCCATGGCACAGGCCAGGCAGCCGGAG
GCTGGGTGGGCCTCAGTGGGGCTGCTGCCAGCCCCAGCACAATAAAATGAAACGTGA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGT
CGACCTGCAGAAGCTTGGCCGCCATGGCCAACTTGTTATTGCAGCTATAATGGTTACAAAT

FIGURE 213

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHGDGVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLAADGTLCPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVIALPLHSLASQALEHGLPDPGSILLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKKDS

Signal sequence:

1-19

Y D E G E G D E E E D D D

FIGURE 214

GCAGGGCAGGTGGGCTCAGGAGGTGCCTCCAGGGGCCAGTGGCCTGAGGCCAGCAAG
GGCTAGGGTCCATCTCCAGTCCCAGGACACAGCAGCGGCCACCATGCCACGCCCTGGCTCC
AGCAGCATCAGAGCAGCCCTGTGGTGGCAGCAAAGTTAGCTGGCTGGCCCTGTGA
GGGGCTTCGCGCTACGCCCTGCGGTGTCCCGAGGGCTGAGGTCTCCTCATCTCCCTAGC
AGTGGATGAGCAACCCAACGGGGCCGGGAGGGAACTGGCCCCGAGGGAGAGGAACCC
AAAGCCACATCTGTAGCCAGGATGAGCAGTGTGAATCCAGGCAGCCCCCAGGACCGGGAGG
CACAGGTGGCCCCCACCACCCGGAGGAGCAGCTCCTGCCCTGTCCGGGGATGACTGATT
TCCTCCGCCAGGCCACCCAGAGGAGAACGCCACCCCGCTGGAGGCACAGGCCATGAGGGC
TCTCAGGAGGTGCTGCTGATGTGGCTCTGGTGTGGCAGTGGCGGCACAGAGCACGCC
CCGGCCCGGCGTAGGGTGTGCTGTCGGGCTACGGGACCCCTGTCTCCAGTCGTTCG
TGCAGCGTGTGTACCAAGCCCTCCTCACCAACCTGCAGGGCACCGGCCCTGCAGCACCTAC
CGAACCATCTATAGGACCGCCTACGCCAGCCCTGGGCTGGCCAGGCCCTGCCAGGCC
CGCGTGTGCCCGGCTGGAAAGAGGACCAGCGGGCTTCCTGGGCTGTGGAGCAGCAATAT
GCCAGCCGCCATGCCGGAACGGAGGGAGCTGTGTCAGCCTGGCGCTGCCGCTGCC
GGATGGCGGGGTGACACTTGCAGTCAGATGTGGATGAATGCAGTGCTAGGAGGGCGGCTG
TCCCCAGCGCTGCATCAACACCGCCGGCAGTTACTGGTGCAGTGTGGAGGGCACAGCC
TGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGGCCCCCAGGGTGGCCCCAACCCG
ACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAGGCTGCAGTCAGGTGGACCTGCT
GGAGGAGAAGCTGCAGCTGGTGCCTGGCCCACTGCACAGCCTGGCTCGCAGGCACTGGAGC
ATGGGCTCCGGACCCGGCAGCCTGGTGCACTCCTCCAGCAGCTGGCGCATCGAC
TCCCTGAGCGAGCAGATTCTCCTGGAGGAGCAGCTGGGTCTGCTCCTGCAAGAAAGA
CTCGTGACTGCCAGCGCTCCAGGCTGGACTGAGCCCTCACGCCGCCCTGCAGCCCCCATG
CCCCTGCCAACATGCTGGGGTCCAGAACGCCACCTGGGTGACTGAGCGGAAGGCCAGGC
AGGGCCTTCCTCCTCTCCCTGGGAGGCTCCCCAGACCCCTGGCATGGGAT
GGGCTGGGATCTCTCTGTGAATCCACCCCTGGCTACCCCAACCCCTGGCTACCCCAACGGCA
TCCCAAGGCCAGGTGGACCCCTCAGCTGAGGGAAAGGTACGAGCTCCCTGCTGGAGCCTGGGAC
CCATGGCACAGGCCAGGCAGCCGGAGGCTGGTGGGCCTCAGTGGGGCTGCTGCC
CCCCAGCACAATAAAATGAAACGTG

FIGURE 215

MRGSQEVLLMWLLVLAVGGTEHAYRPGRRVCAVRAHDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGWKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTCQSDVDECSARRGGCPQRCINTAGSYWCQCWEGHSLAADGTLCPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPLHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSCKDS

Signal sequence:

1-19

FIGURE 216

CCACCGCGTCCGAAGCTGGCCCTGCACGGCTGCAAGGGAGGCTCCTGGACAGGCCAGGCA
GGTGGGCCTCAGGAGGTGCCTCCAGGCCAGTGGCCTGAGGCCAGCAAGGGCTAGGG
TCCATCTCCAGTCCCAGGACACAGCAGGCCACCATGCCACGCCCTGGCTCCAGCAGCAT
CAGCAGCCCCCAGGACCGGGGAGGCACAGGTGGCCCCACCACCCGGAGGAGCAGCTCCTGC
CCCTGTCCGGGGATGACTGATTCTCCTCCGCCAGGCCACCCAGAGGAGAAGGCCACCCGC
CTGGAGGCACAGGCCATGAGGGGCTCTCAGGAGGTGCTGCTGATGTGGCTTCTGGTGGC
AGTGGGCGGCACAGAGCACGCCTACCGGCCGGCGTAGGGTGTGCTGTCGGCTCACG
GGGACCCCTGTCTCGAGTCGTCGTGCAGCGTGTGTACCAAGCCCTCCTCACCACTGCGAC
GGGCACCGGGCTGCAGCACCTACCGAACCATCTATAGGACCGCCTACGCCAGCCCTGG
GCTGGCCCTGCCAGGCCCGCTACGCGTGTGCCGGCTGGAAGAGGAGCAGCGGGCTTC
CTGGGCCTGTGGAGCAGCAATATGCCAGGCCATGCCGGAACGGAGGGAGCTGTGTCAG
CCTGGCCGCTGCCGCTGCCCTGCAGGATGGCGGGTGACACTTGCCAGTCAGATGTGGATGA
ATGCAGTGCTAGGAGGGCGGCTGTCCCCAGCGCTGCGTCAACACCGCCGGCAGTTACTGGT
GCCAGTGTGGAGGGCACAGCCTGTCTGCAGACGGTACACTCTGTGTGCCAAGGGAGGG
CCCCCAGGGTGGCCCCAACCCGACAGGAGTGGACAGTGCAATGAAGGAAGAAGTGCAGAG
GCTGCAGTCCAGGGTGGACCTGCTGGAGGAGAAGCTGCAGCTGGTGTGCCCAAGGGAGGG
GCCTGGCCTCGCAGGCACTGGAGCATGGCTCCGGACCCCGGCAGCCTCCTGGTGCACTCC
TTCCAGCAGCTGGCCGATCGACTCCCTGAGCGAGCAGATTCTCCTGGAGGAGCAGCT
GGGTCTGCTCTGCAAGAAAGACTCGTGACTGCCAGGCCAGGCTGGACTGAGCCCC
TCACGCCGCCCCCTGCAGCCCCCATGCCCTGCCAACATGCTGGGGTCCAGAACCCACCTCG
GGGTGACTGAGCGGAAGGCCAGGCAGGGCTTCCTCCTCTCCCTCCTCGGGAG
GCTCCCCAGACCCCTGGCATGGATGGATGGCTGGATCTTCTCTGTGAATCCACCCCTGGCTACC
CCCACCCCTGGCTACCCCAACGGCATCCAAAGGCCAGGTGGCCCTCAGCTGAGGGAAAGGTAC
GAGCTCCCTGCTGGAGCCTGGACCCATGGCACAGGCCAGGCAGCCGGAGGCTGGGTGGGG
CCTCAGTGGGGCTGCTGCCCTGACCCCCAGCACAATAAAATGAAACGTG

FIGURE 217

MRGSQEVLMMWLLVAVGGTEHAYRPGRRVCARAHGDPVSESFVQRVYQPFLTTCDGHRAC
STYRTIYRTAYRRSPGLAPARPRYACCPGKRTSGLPGACGAAICQPPCRNGGSCVQPGRCR
CPAGWRGDTQSDVDECSARRGGCPQRCVNTAGSYWCQCWEHSLSADGTLCPKGPPRVA
PNPTGVDSAMKEEVQRLQSRVDLLEEKLQLVLAPlHSLASQALEHGLPDPGSLLVHSFQQLG
RIDSLSEQISFLEEQLGSCSKKDS

Signal sequence:

1-19

FIGURE 218

GGTTGCCACAGCTGGTTAGGGCCCCGACCACTGGGGCCCTTGTCAAGGAGGAGACAGCCTCCCGGCCCCGGAG
GACAAGTCGCTGCCACCTTGGCTGCCGACGTGATTCCCTGGACGGTCCGTTCTGCCGTAGCTGCCGCCG
AGTTGGGTCTCCGTGTTCAAGGCCGGCTCCCCCTTCTGGTCTCCCTCTCCGCTGGGCCGGTTATCGGGAGG
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTGATGATGTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT
TTCAAAGCTGGGCTCAGCCTCTGTTCTCTCGTGTAAATCGAAAACCAATTGGAGGAGGAATTCCAATCATGT
CTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAAACTCCAGGCAGGAACACCTTTGCTGTG
ATGGCCCGTGTATGATGGCCCGCAAAAGGGCATTTCTACCTGACCCCTTCTCATCCTGGGACATGTACAC
TCTTCTTCGCCTTGAGTGCCGCTACCTGGCTGAGCTGTCCTGCCATCCCTGTATTGCTGCCATGCTCT
TCCCTTCTCCATGGCTACACTGTTGAGGACCAAGCTTCACTGACCCCTGGAGTGATTCTCGGGCCTACCAAGATG
AAGCAGCTTCATAGAAAATGGAGATAGAAGCTACCAATGGTGCAGGTCAGGCCAGCAGCACCGCCTCGTA
TCAAGAAATTCCAGATAAAACAACCAGATGTGAAACTGAAATACTGTTACACATGCAAGATCTTCCGGCCTCCCC
GGGCCTCCATTGCAAGCATCTGTGACAACCTGTGTGGAGCGCTTCGACCATCACTGCCCTGGTGGGAATTGTG
TTGGAAAGAGGAACCTACCGTACTTCTACCTCTCATCCTTCTCTCCCTCCATCAAATCTATGTCTTCGCT
TCAACATCGTCTATGTGGCCCTCAAATCTTGAAAATTGGCTTCTGGAGACATTGAAAGAAACTCCTGGAACTG
TTCTAGAAGTCCTCATTGCTCTTACACTCTGGCCGTGAGGACTGACTGGATTCTACTTTCTCGTGG
CTCTCAACCAGACAACCAATGAAGAGACATCAAAGGATCATGGACAGGGAAAGATCGCTCCAGAATCCCTACAGCC
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGGCCCTTGCCTGGAGGACTGTGCTGGATCGAAGGGGTA
TTTGCCACTGGAGGAAGTGAAGTCGACCTCCAGTACTCAAGAGACAGTAGCAGCCTTGCACAGAGCC
CAGCCCCACAGAACACCTGAACCAAATGAGATGCCGAGGACAGCAGCACTCCGAAGAGATGCCACCTCCAG
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAGTAGCCTTATCTATGAAAGAGACTTTGTTGTGTT
TAATTAGGGCTATGAGAGATTCAAGTGAGAAGTTAACCTGAGACAGAGAGCAAGTAAGCTGCCCTTTAACT
GTTTTCTTGGCTTGTACCCAGTGCACACTGGCATTCTTGCTGCAAGCTTTAAATTCTGAACCT
CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGCTCTTGGCCCTGGCACTGGTTCT
CCATGGCCTCAGCCACAGGGTCCCCCTTGACCCCTCTCCAGATCCCAGCCTCTGCTTGGGTAC
TGGTCTCATTCTGGGCTAAAAGTTTGAGACTGGCTCAAATCCTCCAAAGCTGCTGCACTGCTGAGTCAGA
GGCAGTCACAGAGACCTCTGGCCAGGGGATCTAACCTGGGTTCTGGGTCTTCAGGACTGAAGAGGAGGAGAG
TGGGGTCAGAAGATTCTCTGCCACCAAGTGCAGTCCAGCATTGCCACAAATCTTTAGGAATGGACAGGTACCT
TCCACTTGTGTANNNNNNNNNNNNNNNNNNNNNNNNNNNTGTTTCTCTGACTCTGCTCCATTAGGAG
CAGGAATGGCAGTAATAAAAGTCTGCACTTGGTCATTCTCTCAGAGGAAGGCCGAGTGTCACTTAAAC
ACTATCCCTCAGACTCCCTGTGTGAGGCCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAG
GCTCTCCTCTCCTCTCCCTCCCCGATGTACCCCTCAAAAAAAAAAATGCTAACAGCTTCCATTAGCCT
CGGCTGAGTGGGAAAGGCCAGCACTGCTGCCCTCTGGTAACTCACCTAAGGCTCGGCCACCTCTGGCT
ATGGTAACCACACTGGGGCTTCCCTCAAGCCCGCTCTCCAGCAGTCCACCGGCAGAGTCCCAGAGCCACTT
CACCCCTGGGGTGGCTGTGGCCCCAGTCAGCTGCTCAGGACCTGCTTATTCAGGAAGAAGATTATGT
ATTATATGTGGCTATATTCTAGAGCACCTGTGTTTCTTCTAAGCCAGGTCTGTCTGGATGACTTAT
GCGGTGGGGAGTGTAAACCGAACCTTCATCTATTGAAGGCATTAAACTGTGCTAATGCA

FIGURE 219

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTFLFLILGTCTLFFAFECRYLAV
QLSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPLALPDEAAFIEMEIEATNGAVPQGQRPPP
RIKNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHCPWVGNCVGKRNRYRYFYL
FILSLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLIKFFTTLWSVVGLTGFHTF
LVALNQTTNEDIKGWSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGR
PPSTQETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMPPPEPPPEPQEAEEAEK

Putative transmembrane domains:

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

FIGURE 220

AAAACCTGTATTTTACAATGCAAATAGACAATNANCCTGGAGGTCTTGAATTAGGTAT
TATAGGGATGGTGGGTTGATTTTNTTCCTGGAGGCTTGGCTTGGACTCTCNCCTTCT
CCCACAGAGCNCTTCGACCATCACTGCCCTGGTGGGAATTGTGTGGAAAGAGGAACTA
CCGCTANTTCTACCTCTTCATCCTTNTCTCTCCNCCTCACAAATCTATGTCTCGCCTCA
ACATCGT

FIGURE 221

GTGTCCTCAGAAAACAGTGGATTAAATCTCCTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAACCGAACCTGACAAAAAGAAGAAAAGAAGAAGA
AAAAAAATCATAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGCAATCTTCAC
GGGCTGGCTGCTCTGTGTCTCTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGAGAGGCCACCTCAGGTGCACTATT
GACAACCGGGTACCCGGTGGCTGGCTAAACCGCAGCACCATCCTCTATGCTGGAAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGCTTCTGAGCAACACCCAAACGCACTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATTC
TTCAGATATCTCCATTAATGAAGGGAACAAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTGGAGACACATCTCTCCAAAGCGGTTGGCTTGTGAGTGAAGAC
GAATACTGGAAATTCAAGGCATCACCGGGAGCAGTCAGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCGTGGTACGGAGAGTAAAGGTACCGTGAACATCCACCATACA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGACAAAAGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTCCCTCTCAAAACTCATCTTCAATGTCTG
AACATGACTATGGGAACACTACACTTGCCTGGCCCTCAACAGCTGGCCACACCAATGCCAGC
ATCATGCTATTGGTCCAGGCGCCGTCAAGCGAGGTGAGCAACGGCACGTCGAGGGAGGGCAGG
CTGCCTGGCTGCCTCTCTGGTCTTGCACCTGCTCTCAAATTTGATGTGAGTGC
ACTTCCCCACCCGGAAAGGCTGCCACCACCAACACAAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTGAGGGAGGGAAACAAAGAATACTTGGGGAAAAGAGTTTTAAAAAGAAATTGAA
AATTGCCTTGCAGATATTAGGTACAATGGAGTTTCTTCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCACCTCTTGGTGCCAGTGTGGCAA
GGGCTCAGCCTCTGCCACAGAGTGCCAACGTGGAACATTCTGGAGCTGCCATCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACACCTCCGGCCAAGCGTGGCGCTCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGAAACGTGAAATAAAAGAGCAAAA
AAAAA

400170832 4002104

FIGURE 222

MKTIQPKMHNSISWAIFTGLAALCLFQGVPRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLSSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAvgFVSEDEYL
EIQGITREQSGDYECSASNDVAAPVVRVVKVTVNYPYISEAKGTGPVGQKGLQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGNYTCVASNKLGHTNASIML
FGPGAVSEVSNGTTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

FIGURE 223

GAAAAAAAATCATGAAAACCATCCAGCCAAAATGCACAATTCTATCTCTGGGCAATCTC
ACGGGGCTGGCTGCTCTGTCTCTTCCAAGGAGTGCCCGTGCAGCGGAGATGCCACCTT
CCCCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCCTCAGGTGCACTA
TTGACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAAT
GACAAGTGGTGCCTGGATCCTCGCGTGGTCCTCTGAGCAACACCCAAACGCAGTACAGCAT
CGAGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACA
ACCACCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCAAAATTGTAGAGATT
TCTTCAGATATCTCCATTAATGAAGGAAACAATATTAGCCTCACCTGCATAGCAACTGGTAG
ACCAGAG

FIGURE 224

ATGGCTGGTGACGGCGGGGCGGGCAGGGGACCGGGCGGGGGAGCGGGCAGCTGCCGGAGCCCTGA
ATCACCGCCTGGCCGACTCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCACTGCTGGCT
GCAGTGCCTCTGGGCTGCCTTGTGGCCCTAGGGGTCAGTACACAGAGACCCATCCCACAGCACCTGCCTTACA
GAGGCCTGCATTGAGTGGCTGGAAAAAAATCCTGGAGTCCCTGGAGGGGTGAGCCCTGTGAGGAACCTTAC
CAGTTCTCCTGTGGGGCTGGATTCGAGGAACCCCTGCCGATGGGCTTCTCGCTGGAACACCTTCAACAGC
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAACACACCACCTCAACTCCAGCAGTGAAGCTGAG
CAGAAGACACAGCGCTTCTACCTATCTTGCCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGGCCACTGAGA
GACCTCATTGAGAAGATTGGTGGTGGAACATTACGGGCCCTGGGACCAAGGACAACCTTATGGAGGTGTTGAAG
GCAGTAGCAGGGACTACAGGGCCACCCATTCTCACCGTCTACATCAGTGCCACTCTAAGAGTTCAAACAGC
AATGTTATCCAGGGGACCACTGCTGGCTTCTGCCCTCTGGGATTACTACTAAACAGAACTGCCAATGAG
AAAGTGTCACTGCTATCTGGATTACATGGAGGAACGGGGATGCTGCTGGGTGGGGCCACCTCCACGAGG
GAGCAGATGCAGCAGGTGCTGGAGATAACAGCTGGCCAACATCACAGTGCCCCAGGACCAAGCGCGAC
GAGGAGAAGATCTACCACAAGATGACATTCGGAGCTGCAGGCTCTGGGCCCTCCATGGACTGGCTTGAGTTC
CTGTCTTCTTGCTGTACCATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTCAG
CAGGTGTAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTCAAAG
ACAACCTCAAGCCTGGACCGACGCTTGAGTCTGACAAGAGAAGCTGCTGGAGACCCCTATGGCACTAAGAAG
TCCGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTGGCTTGCTTGACTCTTC
GTGAAGGCCACGTTGACGGCAAAGAAAGAAATTGAGGAGCTGGGATGATCAGCGAAATCCGGACCGCATTGAG
GAGGCCCTGGGACAGCTGGTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT
GATATGATTGGTTTCCCAGACTTATCCTGGAGGCCAAAGAGCTGGATGATGTTATGACGGGTACGAAATTCT
GAAGATTCTTCTTCCAAAACATGTTGAAATTGTAACACTTCTGCAAGGTTATGCTGACCCAGCTCCGAAG
CCTCCCGAGACCAGTGGAGCATGACCCCCCAGACAGTGGGAATGCCACTACCTTCAACTAAGAATGAGATC
GTCTTCCCGCTGGCATCTGCAGGCCCTTCTATGCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATT
GGTGTGGTATGGGCCATGAGTTGACGCATGCCCTTGATGACCAAGGGCGGAGTATGACAAAGAAGGGAACCTG
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGAACACGGCCTGCAATGGAGGAACAGTACAATCAA
TACCAAGGTCATGGGAGAGGTCAACGGCCAGACGCTGGGGAGAAACATTACTGACAACGGGGGCTGAAG
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGAGGAGCAGCAACTGCCAGCCGTGGGCTCACC
AACCACCAAGCTCTCTCGTGGATTGCCAGGTGTGGTCTGGTCCGCACACCAAGAGAGCTCTCACGAGGGG
CTGGTACCGACCCCCACAGCCCTGCCGCTTCCGCGTGTGGCACTCTCTCAACTCCGTGACTTCTGCGG
CACTTCGGCTGCCCTGTGGCTCCCCATGAACCCAGGGCAGCTGTGAGGTGTGGTAGACCTGGATCAGGGGA
GAAATGGCCAGCTGTCACCAAGACCTGGGCAGCTCTCTGACAAAGCTGTTGCTCTGGGTTGGGAGGAAGCAA
ATGCAAGCTGGCTGGCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCCTCTCAATCACCACATTG
TGCCTCTGCTTTGGGGTGCCCTGCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCACCC
GCCTGGAGAGGTCTGGTGGGAGGCCAGTCCCATAGGAAGGAGTCTGCC

FIGURE 225

MNVALQELGAGSNVGQKGTRQLLGSRTQLELVLAGASLLLALLGCLVALGVQYHRDPSH
STCLTEACIRVAGKILESLDRGVSPCEDFYQFSCGGWIRRNPPLDGRSRWNTFNSLWDQNQA
ILKHLLENTTFNSSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNIITGPWDQDN
FMEVLKAVAGTYRATPFFTVDYIISADSKSSNSNVIQVDQSGLFLPSRDYYLNRTANEKVLTAY
LDYMEELGMLLGGRPTSTREQMQVLEIQLANITVPQDQRRDEEKIYHKMSISELQALAP
SMDWLEFLSFLSPLELSDSEPVVVGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSL
DRRFESAQEKLLETLYGKKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMI
SEIRTAFFEEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSF
FQNMLNLYNFSAKVMADQLRKPPSRDQWSMTPQTVNAYYLPTKNEIVFPAGILQAPFYARNH
PKALNFGGIGVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQV
NGERLNGRQTLGENITDNGGLKAAYNAYKAWLRLKHGEQQLPAVGLTNHQLFFVGFAQVWCS
VRTPESSHEGLVTDPHSPARFRVLGTLNSRDFLRHFGCPVGSPMNPGQLCEVW

Type II Transmembrane domain:

amino acids 32-57

FIGURE 226

FIGURE 227

GGCGGAGGGGGGTGCTGGCGGCCGTGATGGCTGGTACGGCGGGCCGGCAGGGGA
CCGGGGCCGCGGCCGGAGCGGGCAGCTGCCGGAGCCCTGAATCACCGCCTGGCCCGAC
TCCACCATGAACGTCGCGCTGCAGGAGCTGGAGCTGGCAGCAACGTGGGATTCCAGAAGGG
GACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGCTTAGCAGGTGCCTCTAC
TGCTGGCTGCACTGTTCTGGCTGCCTGTGGCCCTAGGGTCCAGTACCAACAGAGACCA
TCCCACAGCACCTGCCTTACAGAGGCCTGCATTGAGTGGCTGGAAAAATCCTGGAGTCCCT
GGACCGAGGGGTGAGCCCTGTGAGGACTTTACCAGTTCTCCTGTGGGGCTGGATTCGGA
GGAACCCCTGCCGATGGCGTCTCGCTGGAACACCTCAACAGCCTCTGGACCAAAAC
CAGGCCATACTGAAGCACCTGCTGAAAACACCACCTCAACTCCAGCAGTGAAGCTGAGCA
GAAGACACAGCGCTTCTACCTATCTGCCTACAGGTGGAGGCATTGAGGAGCTGGAGCCC
AGCCACTGAGAGACCTCATTGAGAAGATTGGTGGTGGAACATTACGGGCCCTGGACAG
GACAACTTATGGAGGTGTTGAAGGCAGTAGCAGGGACCTACAGGGCCACCCATTCTCAC
CGTCTACATCAGTGCCGACTCTAACAGAGTTCCAACAGCAATGTTATCCAGGTGGACCAGTCTG
GGCTCTTCTGCCCTCTGGGATTACTACTTAAACAGAACTGCCATGAGAAAGTAAGGAAC
ATCTTCCGAACCCCATCCCTACCCCTGGCTGAGCTGGCTGATCCCTGTTGACTTTCCCT
TTGCCAAGGGTCAGAGCAGGAAGGTGAGCCTATCCTGTACCTAGTGAACAAACTGCCCT
CCTTCTTCTTCTTCTCCCTCCCTCCCTTCTCCCTTTCTCCCTTCCCT
TCTTATTCTTAGGTTCATAGACACCTACTGTGTGCCAGGTCCAGTGGGGAAATTG
GAGATATAAGTTCCGAGCCATTGCCACAGGAAGCGTTAGTGTGATGGTTCATGGACCT
AGATAGGCTGATAACAAAGCTACAAGAGGGCCTGAGGATTCAAGGAGAGACTTATGGAGCC
AGCAAAGTCTCCTGAAGAGATTGCATTTGAGCCAGGTCCCTGTAG

FIGURE 228

ATGCCTACTACCTTCCA ACTAAGAATGAGATCGTCTTCCCCGCTGGCATCCTGCAGGCC
TTCTATGCCCGCAACCACCCCAAGGCCCTGAACTTCGGTGGCATCGGTGTGGTCATGGCCA
TGAGTTGACGCATGCCTTGATGACCAAGGGCGCGAGTATGACAAAGAACGGGAACTGCGGC
CCTGGTGGCAGAATGAGTCCCTGGCAGCCTCCGGAACCACACGGCCTGCATGGAGGAACAG
TACAATCAATACCAGGTCAATGGGGAGAGGCTCAACGGCGCCAGACGCTGGGGGAGAACAT
TGCTGACAACGGGGGCTGAAGGCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATG
GGGAGGAGCAGCAACTGCCAGCCGTGGGCTCACCAACCACCAGCTCTCTCGTGGGATTT
GCCCAGGTGTGGTGCCTCGTCCGACACCAACAGAGAGCTCTCACGAGGGCTGGTACCGACCC
CCACAGCCCTGCCCGCTTCCGCGTGGCACTCTCTCCAACCTCCGTGACTTCCTGCGGC
ACTTCGGCTGCCCTGTCGGCTCCCCATGAACCCAGGGCAGCTGTGAGGTGTGGTAGACC
TGGATCAGGGGAGAAATGCCAGCTGTCACCAGACCTGGGGCAGCTCTCCTGACAAAGCTGT
TTGCTCTGGTTGGGAGGAAGCAAATGCAAGCTGGCTGGGTCTAGTCCCTCCCCCACA
GGTGACATGAGTACAGACCCCTCCTCAATCACCAATTGTGCCTCTGCTTGGGGTGCCCT
GCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTCACCCCTGCCTGGAAGAG
GTCTGGTGGGAGGCCAGTCCCATAGGAAGGAGTCTGCCTCTGTCCCCAGGCTCACT
CAGCCTGGCGGCCATGGGCCTGCCGTGCCTGCCCCACTGTGACCCACAGGCCTGGTGGTG
TACCTCCTGGACTTCTCCCCAGGCTCACTCAGTGCACCTAGGGTGGACTCAGCTCTGTC
TGGCTCACCCCTCACGGCTACCCCCACCTCACCCGTGCTCCTGTGCCACTGCTCCAGTG
CTGCTGCTGACCTCACTGACAGCTCTAGTGGAAAGCCAAGGGCCTCTGAAAGCCTCCTGC
TGCCCACGTGTTCCCTGGCTGAGAGGGGAAGTGCATATGTGTAGCGGGTACTGGTTCTGT
GTCTTAGGGCACAGCCTAGCAAATGATTGATTCTCCCTGGACAAAGCAGGAAAGCAGATA
GAGCAGGGAAAAGGAAGAACAGAGTTATTTTACAGAAAAGAGGGTGGGAGGGTGTGGTCT
TGGCCCTTATAGGACC

FIGURE 229

CCCCACCGCGTCCGAGCCGCCCGAGAATTAGACACACTCCGGACGCGGCCAAAGCAACCGAGA
GGAGGGGAGGCACAAACACCGAAAAACAAAAGAGAGAAACAACACCAACAACTGGGGTGG
GGGGAAAGAAAGAAAAGAAACCCACCCACCCACCAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAATCCTGTGGCGCCGCGCTGGTCTCCGGGAAGACTCGCCAGCACCAGGGGG
TGGGGGAGTGCAGAGCTGAAAGCTGCTGGAGAGTGAAGCAGCCCTAGCAGGGATGGACATGATG
CTGTTGGTGCAGGGTGCCTGTTGCTCGAACCGAGTGGCTGGCGCGGTGCTCCTCAGCCTGTG
CTGCCTGCTACCCCTCGCCTCCGGCTGGACAGAGTGTGGACTTCCCTGGCGGGCCGTGG
ACAAACATGATGGTCAGAAAAGGGACACGGCGGTGCTTAGGTGTTATTGGAAGATGGAGCT
TCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTATTTGCGGGAGGGTGATAAGTGGTCAGT
GGATCCTCGAGTTCAATTCAACATTGAATAAAAGGGACTACAGCCTCCAGATAACAGAATG
TAGATGTGACAGATGATGGCCCATACAGTGTCTGTTCAGACTCAACATAACACCCAGAAC
ATGCAGGTGCATCTAACGTCAAGTTCCTCTAACAGATATATGACATCTCAAATGATATGAC
CGTCAATGAAGGAACCAACGTCACTCTTACTTGTGTTGCCACTGGGAAACCAAGAGCCTTCCA
TTTCTGGCGACACATCTCCCCATCAGCAAAACATTGAAAATGGACAATTGGACATT
TATGGAATTACAAGGGACCAGGCTGGGAATATGAATGCAGTGCAGGAAAATGCTGTGTCATT
CCCAGATGTGAGGAAAGTAAAAGTTGTTGTCACCTTGCTCCTACTATTAGGAAATTAAAT
CTGGCACCGTGACCCCCGGACGCAGTGGCCTGATAAGATGTGAAGGTGCAAGGTGCGGCCT
CCAGCCTTGAATGGTACAAGGAGAGAAGAAGCTTCAATGGCCAACAAGGAATTATT
TCAAAATTAGCACAAGATCCATTCTCACTGTTACCAACGTGACACAGGAGCATTGGCA
ATTATAACCTGTGTGGCTGCCAACAGCTAGGCACAACCAATGCGAGCCTGCCTTAAACCT
CCAAGTACAGCCCAGTATGAAATTACGGGAGCGCTGATGTTCTTCTCTGCTGGTACCT
TGTGTTGACACTGCTCTTACACAGCATATTCTACCTGAAGAATGCCATTCTACAATAAA
TTCAGGACCCATAAAAGGCTTTAAGGATTCTGAAAGTGTGATGGCTGGATCAAATCT
GGTACAGTTGTTAAAGCAGCGTGGGATATAATCAGCAGTGCCTACATGGGATGATGCC
TTCTGTAGAATTGCTATTATGTAATACTTTAATTCTACTCTTTTGATTAGCTACATTA
CCTTGTGAAGCAGTACACATTGCTCTTTAAGACGTGAAAGCTGAAATTACTTTAG
AGGATATTAAATTGTGATTCTGATGTTGTAATCTACAACCTTCAAAAGCATTAGTCATGGT
CTGCTAGGTTGAGGCTGTAGTTACAAAAACGAATTGCAAGTGAATATGTGATTCTTAA
GGCTGCAATACAAGCATTGCTCCCTGTTCAATAAGAGTCATCCACATTACAAAGATG
CATTTTTCTTTGATAAAAAGCAAATAATTGCCTCAGATTATTCTTCAAATA
TAACACATATCTAGATTTCTGCTGCATGATATTAGGTTCAGGAATGAGCCTTGAAT
ATAACTGGCTGTGCAGCTGCTCTCTTCTGTAAGTTAGCATTGGGTGCGCTTACAT
AATAATATTCTCTTGTCTCCAACATAATATAAAATGTTGCTAAATCTACAATTG
AAGTAAAAATAACCAAGAGTGTGATCAAGTTAAACCATACACTATCTTAAGTAACGAAGGAGC
TATTGGACTGAAAAATCTTCTGACTGACAAATGGGTTGAGAATTGCCCCACACT
AACTCAGTTCTGTGATGAGAGACAATTAAATAACAGTATAGTAAATATACCATATGATTTC
TTTAGTTGAGCTAAATGTTAGATCCACCGTGGGAAATCATTCCCTTAAATGACAGCACA
GTCCACTCAAAGGATTGCCTAGCAATACAGCATCTTCTTCACTAGTCCAAGCAAAAAA
TTTTAAGATGATTGTCAGAAAGGGACAAAGTCCTATCACCTAATATTACAAGAGTGGTA
AGCGCTCATCTTATTTATTTGTGGCAGGTATTGACAGTCGACCTGGAGGGTATGGA
TATGGATATGGACGTTCCAGAGACTATAATGGCAGAAACCAAGGGTGGTTATGACCGCTACTC
AGGAGGAAATTACAGAGACAATTATGACAACGTGAAATGAGACATGCACATAATATAGATACA
CAAGGAATAATTCTGATCCAGGATCGTCCTCCAAATGGCTGTATTATAAAGGTTTTGG
AGCTGCACTGAAGCATCTTATTAGTATATCAACCTTGTGTTAAATTGACCTGCCA
AGGTAGCTGAAGACCTTTAGACAGTCCATCTTTTAAATTCTGCTTAA
AGACAAATTATGGGACGTTGTCAAAAAAAAAAAAAAAA

FIGURE 230

MMLLVQGACCSNQWLAAVLLSLCCLLPSCLPAGQSVDFPWAADVNMVRKGDTAVLRCYLED
GASKGAWLNRSSIIFAGGDKWSVDPRVSISTLNKRDSLQIQNVDVTDDGPYTCVQTQHTP
RTMQVHLTVQVPPKIYDISNDMTVNEGTVNLTCLATGKPEPSISWRHISPSAKPFENGQYL
DIYGITRDQAGEYECSAENAVSFDPVRKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGAGV
PPPAFEWYKGEKKLFNGQQGIIIQNFSTRSILTVTNVTQEHFGNYTCVAANKLGTNASLPL
NPPSTAQYGITGSADVLFSCWYLVLTSSFTSIFYLKNAILQ

Important features of the protein:

Signal peptide:

amino acids 1-31

Transmembrane domain:

amino acids 326-345

N-glycosylation sites.

amino acids 71-75, 153-157, 273-277, 284-288, 292-296, 305-309

Casein kinase II phosphorylation site.

amino acids 147-151, 208-212, 224-228

Tyrosine kinase phosphorylation site.

amino acids 178-186

N-myristoylation sites.

amino acids 7-13, 63-70, 67-73, 151-157, 239-245, 291-297,
302-308, 319-325

Myelin P0 protein:

amino acids 92-121

FIGURE 231

AGTGGTTCGATGGGAAGGATCTTCTCCAAGTGGTCCTCTGAGGGGAGCATTCTGCTGG
CTCCAGGACTTGGCCATCTATAAGCTTGGCAATGAGAAATAAGAAAATTCTCAAGGAGGA
CGAGCTCTGAGTGAGACCCAACAAGCTGCTTACCAAATTGCAATGGAGCCTTCGAAA
TCAATGTTCAAAGCCAAGAGGAGAAATGGGTGAACCTCTCCCTAGCTGTGGTGGTCATC
TACCTGATCCTGCTACCGCTGGCCTGGCTGCTGGTCCAAGTTCTGAATCTGCAGGC
GCGGCTCCGGTCCTGGAGATGTATTCCTCAATGACACTCTGGCGCTGAGGACAGCCC
CCTTCTCCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTG
CAAGTCCTGCAGGCCAACTCACCTGGTCCCGTCAGCCATGAGCAGTGCTGCAGCGGG
AGACAACTTCACTCAGAACCCAGGGATGTTCAGAATCAAAGGTGAACAAGGCCAGGTC
TTCAAGGTCAAGGGGCCATGGCATGCCCTGGTCCCCCTGGCCCGCCGGGACCACCTGCT
GAGAAGGGAGCCAAGGGGCTATGGACGAGATGGAGCAACAGGCCCTGGGACCCCAAGG
CCCACCGGGAGTCAGGGAGAGGCAGGCCCTCAAGGACCCAGGGTGCCTCAGGGAAAGCAAG
GAGCCACTGGCACCCCAGGACCCCAAGGAGAGAACGGCAGCAAAGGCATGGGTCTCATT
GGCCAAAAGGGAAACTGGAACTAAGGGAGAGAACGGAGACCTGGTCTCCAGGAAGCAA
AGGGGACAGGGCATGAAAGGAGATGCAGGGTCATGGGCCTCCTGGAGCCAGGGAGTA
AAGGTGACTTCGGGAGGCCAGGCCACCAGGTTGGCTGGTTTCTGGAGCTAAAGGAGAT
CAAGGACAACCTGGACTGCAGGGTGTCCGGGCCCTGGTGCAGTGGACACCCAGGTGC
CAAGGGTGAGCCTGGCAGTGCTGGCTCCCTGGCAGCAGGACTCCAGGGAGCCCGGGA
GTCCAGGAGCCACAGGCCCTGAAAGGAAGCAAAGGGACACAGGACTCAAGGACAGCAAGGA
AGAAAAGGAGAATCAGGAGTTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGAGCCAGG
GCTGGCAGGTCCAAGGGAGCCCTGGACAAGCTGCCAGAAGGGAGACCAGGGAGTGAAAG
GATCTCTGGGAGCAAGGAGTAAAGGGAGAAAAAGGTGAAAGAGGTGAAAACACTAGTGTCC
GTCAGGATTGTCGGCAGTAGTAACCGAGGCCGGCTGAAGTTACTACAGTGGTACCTGGG
GACAATTGCGATGACGAGTGGCAAATTCTGATGCCATTGCTCTGCCGATGCTGGTT
ACTCCAAAGGAAGGCCCTGTACAAAGTGGAGCTGGCACTGGCAGATCTGGCTGGATAAT
GTTCAGTGTGGGCACGGAGAGTACCCGTGGAGCTGCACCAAGAACAGTGGGCCATCA
TGACTGCAGCCACGAGGAGGACGCAGGCGTGGAGTGCAGCGTCTGAACCCGGAAACCCTTCA
CTTCTCTGCTCCGAGGTGTCTGGCTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGT
TCCCTGGGACAACGTGAGCAGCCTCTGGAGAGGGCCATTAAATAAGCTCAACATCATTGA

FIGURE 232

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886

><subunit 1 of 1, 520 aa, 1 stop

><MW: 52658, pI: 9.16, NX(S/T): 3

MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSIAVVVIYLILLTAGAGL
LVVQVNLQARLRVLEMYFLNDTLAAEDSPSFSSLQSAHPGEHLAQGASRLQVLQAQLTWVR
VSHEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRD
GATGPPSGPQGPPGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGKGE
KGDLGLPGSKGDRGMKGDAVGVMGPPGAQGSKGDFGRPGPPGLAGFPAGKDQGQPGLQGVPG
PPGAVGHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDTGLQGQQGRKGESGVPGPA
GVKGEQGSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSRIVGSSNRGR
AEVYYSGTWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWLDNVQCRGTESTLW
SCTKNSWGHHDSCHEEDAGVECSV

Transmembrane domain:

amino acids 47-66 (type II)

N-glycosylation sites.

amino acids 43-47, 83-87, 136-140

Tyrosine kinase phosphorylation site.

amino acids 432-440

N-myristoylation sites.

amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,
400-406, 441-447, 475-481, 490-496, 515-521

Amidation site.

amino acids 360-364

Leucine zipper pattern.

amino acids 56-78

Speract receptor repeat

amino acids 422-471, 488-519

C1q domain proteins.

amino acids 151-184, 301-334, 316-349

FIGURE 233

CCACCGCGTCCGAAGGCAGACAAAGGTTCAATTGTAAGAAGCTCCTCCAGCACCTCCTCT
CTTCTCCTTTGCCAAACTCACCCAGTGAGTGTGAGCATTAAAGAACATCCTCTGCCAAG
ACCAAAAGGAAAGAAGAAAAGGCCAAAGCCAAATGAAACTGATGGTACTTGTTTTCAC
CATTGGGCTAACCTTGCTGCTAGGAGTTCAAGCCATGCCTGCAAATCGCCTCTTGCTACA
GAAAGATACTAAAAGATCACAACTGTCACAACCTCCGGAAAGGAGTAGCTGACACAG
ATTGATGTCAATGTCCAGGATCATTCTGGATGGGAAGGGATGTGAGATGATCTGTTACTG
CAACTCAGCGAATTGCTCTGCTGCCAAAGACGTTTCTTGACCAAGATCTCTTCG
TGATTCTTGCAACAATCAATGAAATCTCATGTATTCTGGAGAACACCATTCTGATTTC
CCACAAACTGCACTACATCAGTATAACTGCATTCTAGTTCTATATAGTGCAATAGAGCAT
AGATTCTATAAATTCTTACTTGTCTAAGACAAGTAATCTGTGTTAAACAAGTAGTAATAAA
AGTTAATTCAATCTAAAAAAAAAAAAAA

FIGURE 234

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52758
<subunit 1 of 1, 98 aa, 1 stop
<MW: 11081, pI: 6.68, NX(S/T): 1
MKLMVLVFTIGLTLGGVQAMPANRLSCYRKILKDHNCHNLPEGVADLTQIDVNVQDHFWDG
KGCEMICYCNFSELLCCPKDVFFGPKISFVIPCNNQ
```

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation site.

amino acids 72-76

Tyrosine kinase phosphorylation site.

amino acids 63-71

FIGURE 235

CCCCACCGGTCCCGGGACCGCGTGGCTGGACCCCAGGTCTGGAGCGAATTCCAGCCTGCAGGG
CTGATAAGCGAGGCATTAGTGAGAGAGACTTACCCCGCCGTGGTGGAGGGC
GCGCAGTAGAGCAGCACAGGCGGGTCCCGGGAGGCCGGCTCTGCTCGCGCCGAGATG
TGGAACTCCTTCACGAAACCGACTCGGCTGTGGCCACCGCGCGCCCGCGCTGGCTGTG
CGCTGGGGCGCTGGTGCCTGGCGGGTGGCTCTTCCTCGGCTCCCTCTCGGGTGGTTA
TAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTGGAT
GAATTGAAAGCTGAGAACATCAAGAAGTTCTACATAATTTACACAGATACCACATTAGC
AGGAACAGAACAAAACCTTCAGCTGCAAAGCAAATTCAATCCCAGTGGAAAGAATTGGCC
TGGATTCTGTTGAGCTAGCTCATTATGATGTCCTGTTGCCTACCCAAATAAGACTCATCCC
AACTACATCTCAATAATTAAATGAAGATGAAATGAGATTTCAACACATCATTATTGAACC
ACCTCCTCCAGGATATGAAAATGTTGGATATTGTAACCTTCAGTGCCTTCTCTCCTC
AAGGAATGCCAGAGGGCGATCTAGTGTATGTAACATGACGAACTGAAGACTTCTTAAA
TTGGAACGGGACATGAAAATCAATTGCTCTGGAAAATTGTAATTGCCAGATATGGAAAGT
TTTCAGAGGAAATAAGGTTAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTACT
CCGACCCCTGCTGACTACTTGCCTGGGTGAAGTCCTATCCAGACGGTTGGAAATCTCCT
GGAGGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCAGGAGACCCCTCACACC
AGGTTACCCAGAACATGAATATGTTATAGCGTGGAAATTGCAAGGGCTGTTGGTCTTCAA
GTATTCTGTTCATCCAATTGGATACTATGATGACAGAACGCTCTAGAAAAAATGGTGGC
TCAGCACCACCAAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCTACAATGTTGGACCTGG
CTTACTGGAAACTTTCTACACAAAAAGTCAGATGCACTCTACCAATGAAGTGA
CGAGAATTACAATGTGATAGGTACTCTCAGAGGAGCAGTGGAAACCAGACAGATATGTCATT
CTGGGAGGTCAACGGGACTCATGGTGTGGTATTGACCCCTCAGAGTGGAGCAGCTGT
TGTTCATGAAATTGTGAGGAGCTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGAA
CAATTGTTGCAAGCTGGATGCAGAAGAATTGGTCTTGGTTCTACTGAGTGGCA
GAGGAGAATTCAAGACTCCTCAAGAGCGTGGCGTGGCTTATATTAAATGCTGACTCATCTAT
AGAAGGAAACTACACTCTGAGAGTTGATTGATCAGCTGATGTCAGCTGGTACACAACC
TAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTGAAGGCAAATCTTTATGAAAGTTGG
ACTAAAAAAAGTCCTCCCCAGAGTTCACTGGCATGCCAGGATAAGCAAATTGGATCTGG
AAATGATTGAGGTGTTCTCAACGACTTGGAAATTGCTCAGGCAGAGCACGGTATACTA
AAAATTGGAAACAAACAAATTCAAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATAT
GAGTTGGTGGAAAAGTTTATGATCCAATGTTAAATATCACCTCACTGTGGCCAGGTTG
AGGAGGGATGGTGTGAGCTAGCCAATTCCATAGTGCTCCCTTTGATTGTCAGATTATG
CTGTAGTTAAGAAAGTATGCTGACAAAATCTACAGTATTCTATGAAACATCCACAGGAA
ATGAAGACATACAGTGTATCATTGATTCACTTTCTGCAGTAAAGAATTTCAGAAAT
TGCTTCCAAGTCAGTGAGAGACTCCAGGACTTGAACAAAGCAACCAATAGTATTAAGAA
TGATGAATGATCAACTCATGTTCTGGAAAGAGCATTATTGATCCATTAGGGTTACCAAGAC
AGGCCTTTTATAGGCATGTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTC
ATTCCCAGGAATTATGATGCTCTGTTGATATTGAAAGCAAAGTGGACCCCTCAAGGCCT
GGGGAGAAGTGAAGAGACAGATTATGTTGCAAGCTTCACTGTGCAGGAGCTGCAGAGACT
TTGAGTGAAGTAGCCTAAGAGGATTTTAGAGAATCCGTATTGAATTGTTGGTATGTCA
CTCAGAAAGAATCGTAATGGGTATATTGATAAATTAAAATTGGTATATTGAAATAAAAGT
TGAATATTATATAA

FIGURE 236

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFL
DELKAENIKKFLHNFTQIPLLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTH
PNYISIINEDGNEIFNTSLFEP PPPGYENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFF
KLERDMKINC SGKIVIARYGKVFRGNKVNAQLAGAKGVILYSDPADYFAPGVKSYPDGWNL
PGGGVQRGNILNLNGAGDPLTPGYPANEYAYRRGIAEAVGLPSIPVHPIGYYDAQKLLEKMG
GSAPPDSSWRGSLKVPYNVGPGBTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYV
ILGGHRDSWVFGGIDPQSGAAVHVIEVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEW
AEENSRLLQERGVAYINADSSIEGNYTLRVDCPLMYSLVHNLTKEKSPDEGFEGKSLYES
WTKKSPSPEFSGMPRISKLGSGNDFEVFFQRLGIASGRARYTKNWETNKFSGYPLYHSVYET
YELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCCRDYAVVLRKYADKIYSISMKHPQ
EMKTYSVSFDSLFSAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAFIDPLGLP
DRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDP SKAWGEVKRQIYVAFTVQAAAE
TLSEVA

Signal sequence:

amino acids 1-40

N-glycosylation sites.

amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340,
459-463, 476-480, 638-642

Tyrosine kinase phosphorylation sites.

amino acids 363-372, 605-613, 606-613, 617-626

N-myristoylation sites.

amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341,
360-366, 427-433, 529-535, 707-713